Draft Comparative Effectiveness Review

Number xx

Imaging Techniques for the Surveillance, Diagnosis, and Staging of Hepatocellular Carcinoma

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. xxx-xx-xxxx

Prepared by:

Investigators:

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by the XXXXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This report may periodically be assessed for the urgency to update. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program website at: www.effectivehealthcare.ahrq.gov. Search on the title of the report.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact [insert program email address].

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new health care technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

We welcome comments on this systematic review. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Richard G. Kronick, Ph.D.
Director, Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Nahed El-Kassar, M.D., Ph.D.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: <Acknowledgments>.

Key Informants

<Name>
<Place>
<City>, <ST>

Technical Expert Panel

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Peer Reviewers

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Imaging Techniques for the Surveillance, Diagnosis, and Staging of Hepatocellular Carcinoma

Structured Abstract

Objectives. Hepatocellular carcinoma (HCC) is the most common primary malignant neoplasm of the liver, and accurate identification, characterization, and staging of HCC are important for guiding treatment and other clinical decisions. A number of imaging modalities are available for surveillance, diagnosis, and staging of HCC. The purpose of this review is to compare the effectiveness of imaging techniques for HCC on test performance, diagnostic thinking, clinical outcomes and harms.

Data sources. Articles were identified from searches (from 1998 to 2013) of electronic databases including Ovid MEDLINE, Scopus, and the Cochrane Libraries. The searches were supplemented by reviewing reference lists and searching clinical trials registries.

Review methods. We used predefined criteria to determine study eligibility. We selected studies of ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) that evaluated test performance for surveillance, diagnosis, or staging of HCC. We also included randomized trials and comparative observational studies on effects of imaging on diagnostic thinking, clinical outcomes, and harms. The risk of bias (quality) of included studies was assessed, data were extracted, and results were summarized quantitatively (through meta-analysis) and qualitatively. Analyses were stratified by imaging type and unit of analysis (patient or HCC lesion). Additional analyses were conducted to evaluate the effects of the reference standard used and study, patient, tumor, and technical characteristics on estimates of test performance.

Results. Of the 4476 citations identified at the title and abstract level, we screened and reviewed 759 full-length articles. A total of 255 studies were included, 251 of which evaluated test performance. Evidence from surveillance settings was limited, but found no difference between US without contrast and CT in sensitivity for HCC. For identification of HCC in nonsurveillance settings, sensitivity was lower for US without contrast than for CT or MRI (difference in sensitivity based on within-study comparisons of 0.11 to 0.22, using HCC lesions as the unit of analysis). Sensitivity of MRI was higher than CT when HCC lesions were the unit of analysis (pooled difference 0.09. 95% CI 0.06 to 12). For diagnosis of HCC in patients with focal liver lesions, we found no clear differences in sensitivity between US with contrast, CT, and MRI. Across imaging modalities and indications for imaging, specificity was generally 0.85 or higher, but specificity was not reported in a number of studies. For identification of metastatic HCC lesions, sensitivity of ^{18}F-fluorodeoxyglucose (FDG) PET for identification of metastatic HCC lesions was 0.82 (95% CI 0.72 to 0.90), but sensitivity of FDG PET for intrahepatic lesions was poor. Limited evidence suggests that imaging strategies involving more than one imaging modality, in which a positive test is defined as typical imaging findings on one or more imaging modalities, is associated with higher sensitivity than a single test, with little effect on specificity.

Across imaging modalities, factors associated with lower estimates of sensitivity included use of explanted liver as the reference standard, use of HCC lesions as the unit of analysis, smaller HCC lesion size, and more well-differentiated HCC lesions. For MRI, hepatic-specific

contrast agents were associated with slightly higher sensitivity than nonspecific contrast agents. For PET, limited evidence suggested higher sensitivity with use of PET/CT than with PET alone and with ^{11}C-acetate than with FDG.

Evidence on the comparative effects of imaging for HCC on diagnostic thinking was extremely limited. The proportion of patients correctly assessed with CT for transplant eligibility based on Milan criteria ranged from 40 percent to 96 percent. Evidence on the effects of surveillance with imaging versus no surveillance on clinical outcomes was limited to a single randomized trial. Although it found an association between surveillance with US and alpha-fetoprotein (AFP) and decreased liver-specific mortality, the trial was conducted in China, potentially limiting applicability to screening in the United States, and there were important methodological shortcomings. Evidence on comparative harms associated with imaging was also extremely limited, but indicate low rates of serious direct harms.

Conclusions. Based on estimates of test performance, several imaging modalities appear to be reasonable options for surveillance, diagnosis, or staging of HCC. Although there are some potential differences in test performance between different imaging modalities and techniques, more research is needed to understand the effects of such differences on diagnostic thinking and clinical outcomes.

Contents

Tables

Figures

Appendices

Executive Summary

Background and Objectives

Hepatocellular carcinoma (HCC) is the most common primary malignant neoplasm of the liver, usually developing in individuals with chronic liver disease or cirrhosis. Worldwide, it is the fifth most common cancer and the third most common cause of cancer death.[1] The National Cancer Institute, attributed 156,940 deaths to liver and intrahepatic bile duct cancer in the United States in 2011, with 221,130 new cases diagnosed.[2]

The American Association for the Study of Liver Diseases (AASLD) recommends surveillance for the following groups at high risk for developing HCC: Asian male HBV carriers over age 40, Asian female HBV carriers over age 50, HBV carriers with a family history of HCC, African/North American black HBV carriers, HBV or HCV carriers with cirrhosis, all individuals with other causes for cirrhosis (including alcoholic cirrhosis), and patients with stage 4 primary biliary cirrhosis.[3]

HCC is an aggressive tumor associated with poor survival without treatment.[4] However, when diagnosed early, HCC may be amenable to potentially curative therapy. The three phases of pretherapy evaluation of HCC include surveillance, diagnosis, and staging.[3] Surveillance is the use of periodic testing to identify lesions in the liver that are clinically suspicious for HCC. The diagnosis phase involves the use of additional tests (radiological and/or histopathological) to confirm that the detected lesion is indeed HCC. Staging determines the extent and severity of a person's cancer to inform prognosis and treatment decisions. A number of staging systems are available, including the widely used TNM (tumor, node, metastasis) staging system and the more recent Barcelona Clinic Liver Cancer (BCLC) staging system,[5] which has become the de facto staging reference standard;[3] the Milan criteria are used to identify patients likely to experience better posttransplantation outcomes, though other methods have been proposed.[6]

Imaging techniques to identify the presence of lesions, diagnose HCC, and determine the stage of the disease include: ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). Understanding the diagnostic accuracy of imaging methods and how they affect clinical decisionmaking and, ultimately, patient outcomes is a challenge. Imaging techniques may be used alone, in various combinations or algorithms, and/or with liver-specific biomarkers, resulting in many potential comparisons. Technical aspects of imaging methods are complex and continuously evolving.

Diagnostic accuracy studies use different reference standards, such as explanted liver specimens from patients undergoing transplantation, percutaneous or surgical biopsy, imaging, clinical followup, or combinations of these methods. Use of these different reference standards introduces heterogeneity that may limit comparisons of techniques. Reference standards also are susceptible to misclassification due to sampling error, inadequate specimens, insufficient followup, or other factors. Other considerations, including risk factors for HCC and lesion characteristics, such as tumor size or degree of differentiation, severity of hepatic fibrosis, and etiology of liver disease, may impact the diagnostic accuracy or clinical utility of imaging strategies.

Accurate identification and staging of HCC is critical for providing optimal patient care. However, clinical uncertainty remains regarding the best imaging strategies. The purpose of this report is to comprehensively review the comparative effectiveness and diagnostic performance of different imaging modalities and strategies for surveillance, diagnosis, and staging of HCC.

Scope and Key Questions

The Key Questions and corresponding analytic frameworks used to guide this report are shown below. Separate analytic frameworks address surveillance (Figure A), diagnosis (Key Figure B), and staging (Figure C). The analytic frameworks show the target populations, interventions (imaging tests), and outcomes (diagnostic accuracy, diagnostic thinking, clinical outcomes, and harms) that we examined.

Key Question 1. What is the comparative effectiveness of available imaging-based surveillance strategies (listed below under interventions for KQ 1), used singly or in sequence for detecting hepatocellular carcinoma (HCC) among individuals undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for HCC)?

a. What is the comparative test performance of imaging-based surveillance strategies for detecting HCC?
    i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
    ii. How is the comparative effectiveness modified by patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, or other factors (e.g., results of biomarker tests, setting)?
b. What is the comparative effectiveness of imaging-based surveillance strategies on intermediate outcomes like diagnostic thinking?
c. What is the comparative effectiveness of imaging-based surveillance strategies on clinical and patient-centered outcomes?
d. What are the adverse effects or harms associated with imaging-based surveillance strategies?

Key Question 2. What is the comparative effectiveness of imaging techniques, used singly, in combination, or in sequence, in diagnosing HCC among individuals in whom an abnormal lesion has been detected while undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for HCC) or through the evolution of symptoms and abdominal imaging done for other indications?

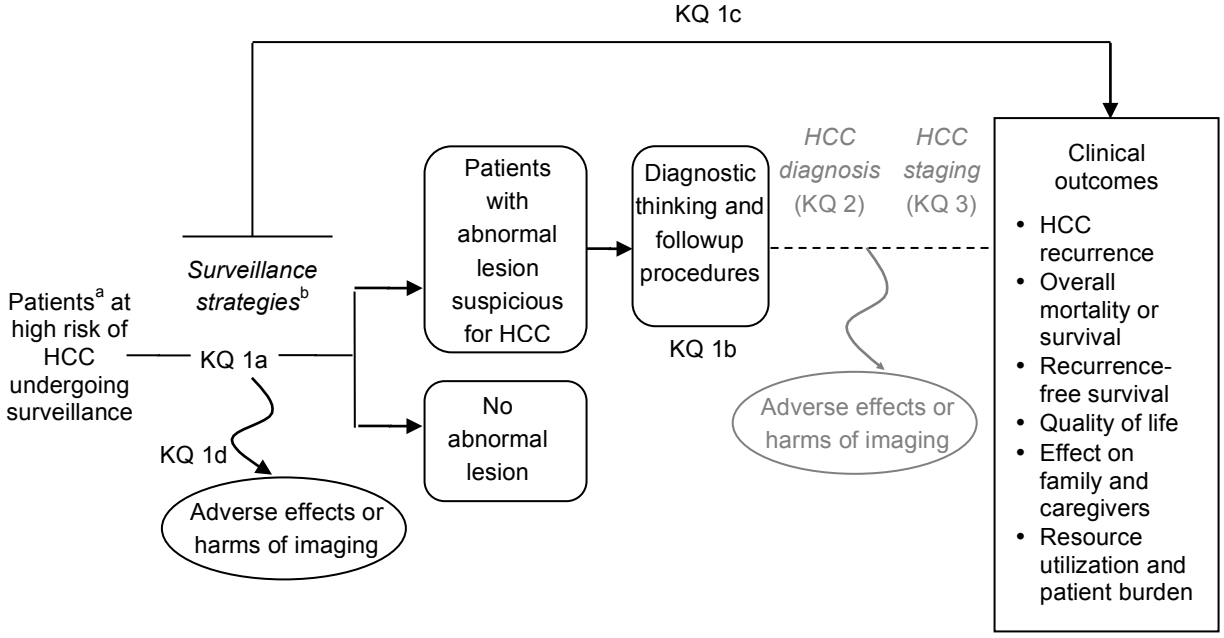
a. What is the comparative test performance of imaging techniques for diagnosing HCC?
    i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical imaging and followup?
    ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?
b. What is the comparative effectiveness of the various imaging techniques on intermediate outcomes such as diagnostic thinking and use of additional diagnostic procedures such as fine-needle or core biopsy?
c. What is the comparative effectiveness of the various imaging techniques on clinical and patient-centered outcomes?

d. What are the adverse effects or harms associated with imaging-based diagnostic strategies?

Key Question 3. What is the comparative effectiveness of imaging techniques, used singly, in combination, or in sequence in staging HCC among patients diagnosed with HCC?

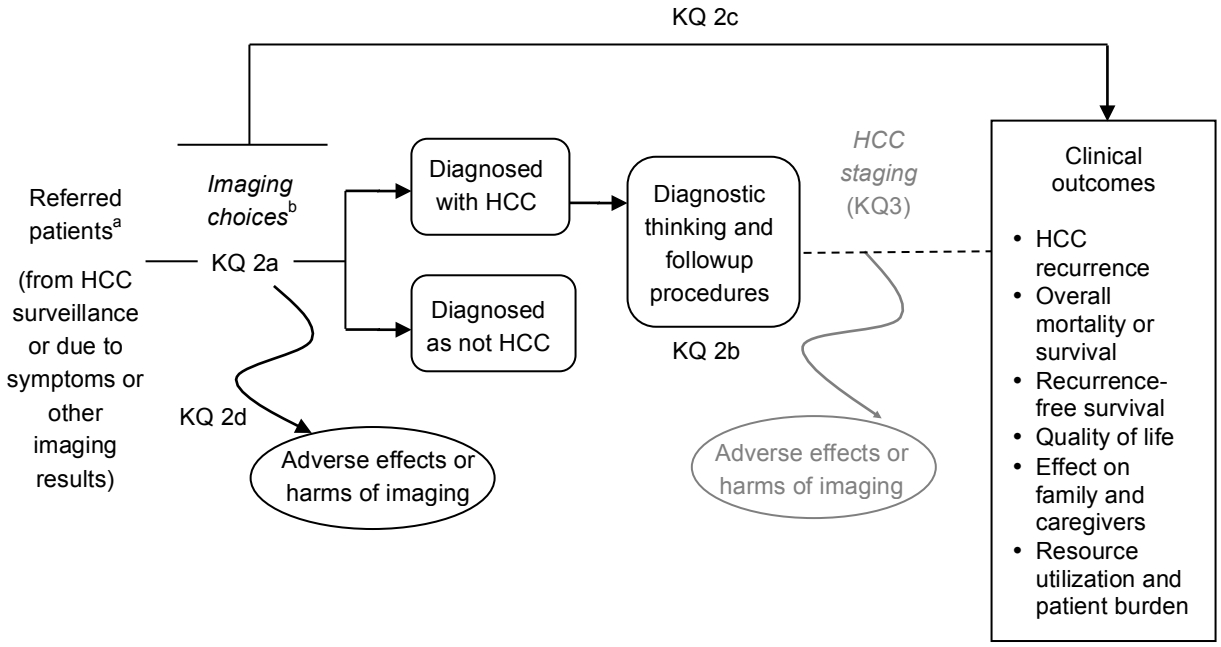
a. What is the comparative test performance of imaging techniques to predict HCC tumor stage?
   i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
   ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?
b. What is the comparative test performance of imaging techniques on diagnostic thinking?
c. What is the comparative effectiveness of imaging techniques on clinical and patient-centered outcomes?
d. What are the adverse effects or harms associated with imaging-based staging strategies?

Figure A. Analytic framework—surveillance (Key Question 1)



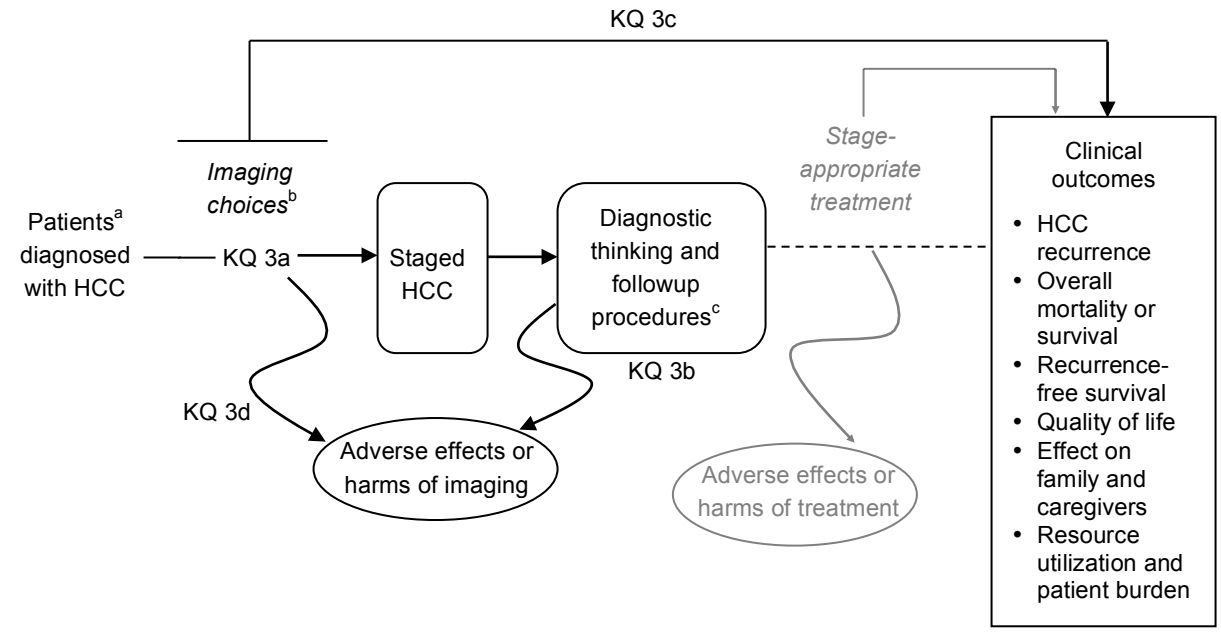
[a] Potential modifiers of test performance include patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, and other factors (e.g., biomarker levels, setting).
[b] Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers
HCC = hepatocellular carcinoma; KQ = Key Question

Figure B. Analytic framework—diagnosis (Key Question 2)



[a] Potential modifiers of test performance include patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, and other factors (e.g., biomarker levels, setting).
[b] Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers
HCC = hepatocellular carcinoma; KQ = Key Question

Figure C. Analytic framework—staging (Key Question 3)



[a] Potential modifiers of test performance include patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, and other factors (e.g., biomarker levels, setting).
[b] Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers
[c] Followup procedures include biopsy.
HCC = hepatocellular carcinoma; KQ = Key Question

Methods

The methods for this systematic review follow the methods suggested in the AHRQ Effective Health Care program methods guides.[7, 8]

Searching for the Evidence

For the primary literature, we searched Ovid MEDLINE®, Scopus, Evidence-Based Medicine Reviews (Ovid), the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, and the Health Technology Assessment Database from 1998 through March 2013. We searched for unpublished studies in clinical trial registries (ClinicalTrials.gov, Current Controlled Trials, ClinicalStudyResults.org, and the WHO International Clinical Trials Registry Platform), regulatory documents (FDA Medical Devices Registration and Listing), and individual product Web sites. Scientific information packets (SIPs) were solicited.[9] We also searched the reference lists of relevant studies and previous systematic reviews for additional studies.

Study Selection

We developed criteria for inclusion and exclusion of studies based on the Key Questions and the populations, interventions, comparators, outcomes, timing, and setting (PICOTS) of interest. Titles and abstracts from all searches were reviewed for inclusion. Full-text articles were obtained for all articles identified as potentially meeting inclusion criteria. Papers were selected for inclusion in our review if they were about imaging for HCC, were relevant to one or more Key Questions, met the predefined inclusion criteria, and reported original data.

We excluded: studies that reported diagnostic accuracy of imaging for non-HCC malignant lesions; studies of nonspiral CT and MRI using machines ≤1.0 T, as these are considered outdated techniques;[10] studies that evaluated MRI with agents that are no longer produced commercially and are unavailable for clinical use; techniques not typically used in the United States for diagnosis and staging of HCC; studies published prior to 1998; studies in which imaging commenced prior to 1995, unless those studies reported use of imaging meeting minimum technical criteria; and studies of intraoperative US. We also excluded studies published only as conference abstracts, non-English language articles, and studies of nonhuman subjects.

For studies of test performance (e.g., sensitivity, specificity, and likelihood ratios), we included studies that evaluated one or more imaging methods against a reference standard. Reference standards were histopathology (based on explanted liver or nonexplant histological specimen from surgery or percutaneous biopsy), imaging and clinical followup, or some combination of these standards. We excluded studies in which the reference standard involved the imaging test under evaluation and studies that had no reference standard (i.e., reported the number of lesions identified with an imaging technique but did not evaluate accuracy against a reference technique).

To assess comparative effects of imaging on clinical outcomes (e.g., mortality, HCC recurrence, quality of life, and harms), we included randomized controlled trials that compared different imaging modalities or strategies. A systematic review funded by the U.S. Department of Veterans Affairs Evidence Synthesis Program on effects of screening for HCC on clinical outcomes is currently in progress that will also include comparative observational studies.[11]

To assess comparative effects of imaging on intermediate outcomes (e.g., effects on diagnostic thinking, clinical thinking, and resource utilization), we included randomized trials and cohort studies that compared different imaging modalities or strategies.

Data Abstraction and Data Management

We extracted the following data from included studies into evidence tables using Excel spreadsheets: study design, year, setting, country, sample size, method of data collection (retrospective or prospective), eligibility criteria, population and clinical characteristics (including age, sex, race, underlying cause of liver disease, proportion of patients in sample with HCC, HCC lesion size, and proportion with cirrhosis), the number of readers, criteria used for a positive test, and the reference standard used. We abstracted results for diagnostic accuracy, intermediate outcomes, and clinical outcomes, including results stratified according to patient, lesion, and imaging characteristics. Technical information for different imaging tests was abstracted.[10]

Assessment of Methodological Risk of Bias of Individual Studies

We assessed risk of bias (quality) for each study based on predefined criteria. Randomized trials and cohort studies were evaluated using criteria and methods developed by the U.S. Preventive Services Task Force.[12] These criteria were applied in conjunction with the approach recommended in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.[7] Studies of diagnostic test performance were assessed using the approach recommended in the AHRQ Methods Guide for Medical Test Reviews,[8] which is based on methods developed by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) group.[13] Individual studies were rated as having "low," "moderate," or "high" risk of bias.

Data Synthesis

We performed meta-analyses on measures of test performance in order to help summarize data and obtain more precise estimates.[14] All quantitative analyses were conducted using SAS 9.3 (SAS institute Inc., Cary, NC). We only pooled studies that were clinically comparable and could provide a meaningful combined estimate (based on the variability among studies in design, patient population, imaging methods, and outcomes) and magnitude of effect size. We conducted separate analyses for each imaging modality, stratified according to the unit of analysis used (patients with HCC, HCC lesions, or liver segments with HCC).

We evaluated a number of potential sources of heterogeneity and modifiers of diagnostic accuracy. We performed analyses stratified according to the reference standard used and on domains related to risk of bias, aspects of study design (retrospective or prospective, use of a confidence rating scale), setting (based on country in which imaging was performed), and technical factors (such as scanner types, type of contrast or tracer used, use of recommended imaging phases, timing of delayed phase imaging, and section thickness). We also evaluated diagnostic accuracy in subgroups stratified according to HCC lesion size, degree of tumor differentiation, and tumor location, as well as patient characteristics such as severity of underlying liver disease, underlying cause of liver disease, and body mass index.

We performed separate analyses on the subset of studies that directly compared two or more imaging modalities or techniques in the same population against a common reference standard. Research indicates that results based on such direct comparisons differ from results based on

noncomparative studies, and may be better suited for evaluating comparative diagnostic test performance.[15]

We did not perform meta-analysis on staging accuracy and intermediate or clinical outcomes due to the small number of studies. Rather, we synthesized these studies qualitatively, using the methods described below for assessing the strength of evidence.

Grading the Strength of Evidence for Individual Comparisons and Outcomes

The strength of evidence for each key question was assessed by one researcher for each outcome described in the PICOTS using the approach described in the AHRQ Methods Guide.[7] The strength of evidence was based on the overall quality of each body of evidence, based on the risk of bias (graded low, moderate, or high); the consistency of results between studies (graded consistent, inconsistent, or unknown/not applicable when only one study was available); the directness of the evidence linking the intervention and health outcomes (graded direct or indirect); and the precision of the estimate of effect, based on the number and size of studies and confidence intervals for the estimates (graded precise or imprecise. We did not assess studies of diagnostic test performance for publication bias using graphical or statistical methods because research indicates that such methods can be very misleading. Rather, we searched for unpublished studies through searches of clinical trials registries and regulatory documents and by soliciting SIPs.

Assessing Applicability

We recorded factors important for understanding the applicability of studies, such as whether the publication adequately described the study population, the country in which the study was conducted, the prevalence of HCC in the patients who underwent imaging, the magnitude of differences in measures of diagnostic accuracy and clinical outcomes, and whether the imaging techniques were reasonably representative of standard practice.[16] We also recorded the funding source and role of the sponsor.

Results

The bulk of the available evidence addresses diagnostic accuracy of different imaging techniques for hepatocellular carcinoma (HCC). Very few studies compared effects of different imaging modalities or strategies on diagnostic thinking and clinical outcomes, and almost no studies reported harms.

Results of Literature Searches

We reviewed titles and abstracts of the 4476 citations identified by literature searches. Of these, 759 articles appeared to meet inclusion criteria and were selected for further full-text review. Of the 759 articles reviewed at the full-text level, a total of 255 studies met inclusion criteria.

We identified 251 studies that evaluated diagnostic accuracy of imaging tests. Of these, 60 evaluated ultrasound imaging, 125 evaluated computed tomography, 117 evaluated magnetic resonance imaging, and 31 evaluated positron emission tomography. Some studies evaluated more than one imaging modality. We rated 3 studies low risk of bias, 162 moderate risk of bias, and 86 high risk of bias. Almost all studies reported sensitivity, but only 117 reported specificity or provided data to calculate specificity. We found 141 studies avoided use of a case-control design, 137 used blinded ascertainment, and 69 used a prospective design. More studies were conducted in Asia (163 studies) than in the United States or Europe (86 studies). In 136 studies, imaging was conducted starting in or after 2003.

Data for outcomes other than measures of test performance were sparse. Seven studies reported comparative effects on diagnostic thinking, three studies reported comparative clinical and patient-centered outcomes, and three studies reported harms associated with imaging for HCC.

Key Question 1. What is the comparative effectiveness of available imaging-based surveillance strategies for detecting HCC among individuals undergoing surveillance for HCC?

Six studies evaluated diagnostic accuracy of imaging techniques for surveillance and 174 studies reported diagnostic accuracy in nonsurveillance settings (e.g., imaging performed to assess detection rates in a series of patients undergoing treatment for HCC or patients with otherwise known prevalence of HCC prior to imaging). Four studies of PET evaluated accuracy specifically for identification of recurrent HCC. One randomized trial (rated high risk of bias) evaluated clinical outcomes associated with imaging-based surveillance versus no screening, and two trials evaluated clinical outcomes associated with different US surveillance intervals. No study compared effects of different imaging surveillance strategies on diagnostic thinking or clinical decisionmaking. Two studies reported harms associated with imaging for HCC. Table A summarizes the key findings and strength of evidence for these studies.

Key Question 2. What is the comparative effectiveness of imaging techniques in diagnosing HCC among individuals in whom an abnormal lesion has been detected while undergoing surveillance for HCC or through the evolution of symptoms and abdominal imaging done for other indications?

Forty-four studies evaluated diagnostic accuracy of imaging tests in diagnosing HCC among individuals in whom an abnormal lesion has been detected and 15 studies evaluated the accuracy

of imaging tests for distinguishing HCC from another specific type of liver lesion. No study compared effects of different imaging modalities or strategies on diagnostic thinking or on clinical or patient-centered outcomes. One study reported harms. Table A summarizes the key findings and strength of evidence for these studies.

Key Question 3. What is the comparative effectiveness of imaging techniques in staging HCC among patients diagnosed with HCC?

Six studies reported test performance of various imaging techniques for staging of patients with HCC based on TNM criteria. Ten studies reported test performance of PET for detection of metastatic disease. Seven studies reported effects of imaging on transplant decisions and one study reported comparative effects of imaging on clinical and patient-centered outcomes. No study reported harms associated with imaging for HCC staging. Table A summarizes the key findings and strength of evidence for these studies.

Discussion

Key Findings and Strength of Evidence

The key findings of this review, including strength of evidence grades, are summarized in Table A. The great preponderance of evidence on imaging for HCC was in the area of diagnostic test performance. However, few studies evaluated test performance of imaging for HCC in true surveillance settings of patients at high risk for HCC, but without a prior diagnosis of HCC, undergoing periodic imaging. Among the limited evidence available in this setting, there was no clear difference between US without contrast and CT, based on across-study comparisons of sensitivity. Two studies that directly compared sensitivity of US without contrast and CT did report lower sensitivity with US, but data are too limited to draw strong conclusions.[17,18]

Many more studies evaluated test performance of imaging for HCC in populations of patients undergoing treatment such as liver transplantation, hepatic resection, or ablation therapy, or in series of patients previously diagnosed with HCC or with HCC and other liver conditions. Such studies were considered as part of Key Question 1 with studies of surveillance because they were not designed to further characterize previously identified HCC lesions (the focus of Key Question 2). Rather, their purpose was to evaluate test performance for lesion identification, therefore providing information that could potentially be extrapolated to surveillance. However, we analyzed these studies separately from studies conducted in true surveillance settings, given the differences in the reason for imaging and the populations evaluated, including a generally much higher prevalence of HCC, with some studies only enrolling patients with HCC. In these studies, sensitivity was lower for US without contrast than for CT or MRI, with a difference based on within-study (direct) comparisons that ranged from 0.11 to 0.22. MRI and CT performed similarly when patients with HCC were the unit of analysis, but sensitivity of MRI was higher than CT when HCC lesions were the unit of analysis (pooled difference 0.09. 95% CI 0.06-12).

Ultrasound without contrast did not perform better than ultrasound with contrast for identification of HCC.[19,20] This is probably related to the short duration in which microbubble contrast is present within the liver, so that it is not possible to perform a comprehensive contrast-enhanced examination of the liver.[21] Rather, the main use of ultrasound with contrast appears to be for evaluation of previously identified focal liver lesions.

For characterization of previously identified lesions, we found no clear differences in sensitivity between US with contrast, CT, and MRI. Although some evidence was available on

the accuracy of imaging modalities for distinguishing between HCC and other (non-HCC) liver lesions, it was not possible to draw strong conclusions due to variability in the types of non-HCC lesions evaluated (e.g., regenerative nodules, dysplastic nodules, hypervascular pseudolesions, hemangiomas, and others), small numbers of studies, and some inconsistency in findings.

Studies of patients with HCC were generally associated with somewhat higher sensitivity than studies that used HCC lesions as the unit of analysis. Studies that used explanted livers as the reference standard reported lower sensitivity than studies that used a nonexplant reference standard. Use of multiple reference standards poses a challenge to assessment of diagnostic accuracy.[22] Across imaging modalities, sensitivity was markedly lower for HCC lesions <2 cm versus those >2 cm (differences in sensitivity ranged from 0.30 to 0.39), and further declined for lesions <10 mm in diameter. Evidence also consistently indicated substantially lower sensitivity for well-differentiated lesions than moderately- or poorly-differentiated lesions.

Evidence on the effects of other patient, tumor, and technical factors on test performance was more limited. For US, there was no clear effect of use of Doppler, lesion depth, or body mass index on test performance. For CT, some evidence indicated higher sensitivity for studies that used a contrast rate of ≥3 ml/s than those with a contrast rate <3 ml/s, and for studies that used delayed phase imaging. For MRI, hepatic-specific contrast agents were associated with slightly higher sensitivity than nonspecific contrast agents, but there were no clear effects of magnetic field strength (3.0 vs. 1.5 T), use of delayed phase imaging, timing of delayed phase imaging ≥120 seconds after administration of contrast of <120 s), section thickness (≤5 mm vs. >5 mm), or use of diffusion-weighted imaging. For identification of intrahepatic HCC lesions, limited evidence found PET with ^{11}C-acetate and other alternative tracers such as ^{18}F-fluorocholine and ^{18}F-fluorothymidine associated with substantially higher sensitivity than FDG PET. Sensitivity of FDG PET was lower than sensitivity of FDG PET/CT.

The limited available evidence suggests that using multiple imaging tests and defining a positive test as typical imaging findings on at least one imaging modality increases sensitivity without substantively reducing specificity.

Conclusions were generally robust on sensitivity and stratified analyses based on study factors such as setting (Asia vs. United States or Europe), prospective collection of data, interpretation of imaging findings blinded to results of the reference standard, avoidance of case-control design, and overall risk of bias.

Across analyses, specificity was generally high, with most pooled estimates around 0.85 or higher, and few clear differences between imaging modalities. However, many studies did not report specificity and pooled estimates of specificity were frequently imprecise, precluding strong conclusions regarding comparative test performance. Since likelihood ratios are sensitive to small changes in estimates when the specificity is high, it was also difficult to draw strong conclusions regarding comparative diagnostic test performance based on differences in positive or negative likelihood ratios. Most likelihood ratio estimates fell into or near the "moderately useful" range (positive likelihood ratio of 5-10 and negative likelihood ratio of 0.1-0.2), with the exception of FDG PET for identification of intrahepatic HCC lesions, which was associated with a negative likelihood ratio of 0.50.

Evidence regarding the accuracy of imaging modalities for staging was primarily limited to CT. Most studies addressed accuracy of CT, with 28 percent to 58 percent correctly staged based on TNM criteria, with somewhat more understaging (25% to 52%) than overstaging (2% to 27%). Studies on the accuracy of imaging for identifying metastatic HCC disease were primarily limited to FDG PET or PET/CT, with a pooled sensitivity of 0.82 to 0.85.

Evidence on the comparative effectiveness of imaging for HCC on diagnostic thinking, use of subsequent procedures, or resource utilization was extremely limited. In studies that compared

the accuracy of transplant decisions based on CT against primarily explanted livers as the reference standard, the proportion correctly assessed for transplant eligibility based on Milan criteria ranged from 40 percent to 96 percent. Evidence on the effects of surveillance with imaging versus no surveillance on clinical outcomes was limited to a single randomized trial.[23]

Evidence on comparative harms associated with imaging was also extremely limited, with no study measuring downstream harms related to false-positive tests or subsequent workup, or potential harms related to labeling or psychological effects. A handful of studies reported low rates of serious direct harms (e.g., allergic reactions) associated with imaging. However, evidence on administration of contrast for radiological procedures in general also suggests a low rate of serious adverse events. No study on US with contrast reported harms. Although PET and CT are associated with risk of radiation exposure, no study of imaging for HCC was designed to evaluate potential long-term clinical outcomes associated with radiation exposure.

Findings in Relationship to What is Already Known

Unlike our review, several previously published reviews on detection of HCC and evaluation of focal liver lesions found no clear differences in test performance between US, CT, and MRI for HCC.[24-27] Several factors may explain these discrepancies—we included more studies than any prior review, separately analyzed studies based on the reason for imaging, stratified studies according to the unit of analysis, and focused on within-study (direct) comparisons of two or more imaging modalities against a common reference standard instead of relying primarily or solely on across-study (indirect) estimates of test performance. Our review is consistent with previous reviews regarding lower sensitivity of imaging for detection of small and well-differentiated HCC lesions.

Our findings regarding test performance of PET for detection of metastatic HCC are consistent with a recently published systematic review and meta-analysis.[28] Like our review, a recent systematic review found insufficient evidence to determine effects of surveillance with imaging on clinical outcomes.[29] A systematic review on screening for HCC in chronic liver disease funded by the U.S. Department of Veterans Affairs is currently in progress.[11]

Applicability

A number of potential issues could impact the applicability of our findings. Over half of the studies were conducted in Asia, where the prevalence, underlying causes, course, evaluation, and management of chronic liver disease may be different than in the United States. To mitigate potential effects of study country on applicability, we excluded invasive imaging techniques not typically used in the United States such as CT arterial portography and CT hepatic arteriography, as well as imaging techniques considered inadequate in the United States (such as C-arm CT). We also performed stratified analyses focusing on studies performed in the United States and Europe to evaluate effects on estimates of diagnostic accuracy and found no clear effects on estimates.

Imaging techniques are rapidly evolving, which is another factor that could affect applicability. To mitigate effects of outdated techniques on applicability, we excluded imaging technologies considered outdated, such as MRI with magnetic field strength <1.5 T and nonspiral CT, and only included studies published since 1998. We also performed additional analyses on technical factors such as contrast rate, imaging phases evaluated, timing of imaging phases, section thickness, use of hepatobiliary contrast (for MRI), use of diffusion-weighted imaging, and newer technologies such as dual-source or spectral CT. We included studies of US with microbubble contrast event though no agent is currently approved for abdominal imaging in the

United States, because efforts to obtain FDA approval are ongoing and this technique is commonly used in other geographic areas of the world, including Canada and Europe.

As noted above, few studies were performed in true surveillance settings, i.e., in patients at high risk for HCC but not previously diagnosed with this condition. Rather, most studies of test performance that were not performed specifically to evaluate or characterize previously identified lesions were conducted in patients undergoing imaging for other reasons, including series of patients undergoing liver transplantation, surgical resection, or other treatments for HCC. Although such studies are likely to provide some useful findings regarding diagnostic accuracy, results may not be directly applicable to patients undergoing surveillance. In particular, the high prevalence of HCC (many studies only enrolled patients with HCC) could overestimate test performance in true surveillance settings, in which the prevalence of HCC would be much lower.[30]

Implications for Clinical and Policy Decisionmaking

Our review has important potential implications for clinical and policy decisionmaking. Due to the lack of direct evidence regarding clinical benefits and downstream harms associated with different imaging tests for surveillance, diagnosis, and staging of HCC, most decisions regarding use of imaging tests must necessarily be made primarily on the basis of diagnostic test performance. Despite limited evidence in true surveillance settings, our study support current recommendations from the AASLD for US without contrast for surveillance of HCC in at-risk populations.[3] Although sensitivity of CT and MRI for identifying HCC was higher than US in studies conducted in nonsurveillance settings, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates.

In patients found to have an HCC lesion on surveillance, our review supports use of CT and MRI to further characterize lesions >1 cm in size, as in the AASLD guideline, based on high sensitivity and specificity. Evidence is very limited but appears consistent with the sequential diagnostic imaging algorithm as outlined in the AASLD guideline, in which typical findings for HCC on sequentially performed CT or MRI are considered sufficient to make a diagnosis.

Our findings also support minimal technical specifications for MRI and CT for HCC imaging as suggested in recent guidance, such as those regarding minimum contrast rates and use of delayed phase imaging.[10] Evidence suggesting superior test performance of MRI with hepatic-specific versus nonhepatic contrast appears promising, though differences were relatively small. Therefore, clinical and policy decisions around use of nonhepatic contrast may be impacted by additional factors other than test performance, such as cost, harms, or convenience.

Although US with contrast was associated with similar test performance as MRI and CT for evaluation of lesions, no microbubble contrast agents are currently approved for use in the United States. Although the role of PET is likely to remain focused on identification of metastatic HCC and staging, additional research could help clarify the role of PET with alternative tracers for identification and evaluation of intrahepatic HCC.

Research Gaps

Significant research gaps limit the full understanding of the comparative effectiveness of imaging for surveillance, diagnosis, and staging of HCC. The only randomized trial of effects of surveillance for HCC with imaging on clinical outcomes had important methodological shortcomings and was performed in China, potentially limiting applicability to screening in the United States[23] Although conducting a randomized trial of surveillance versus no screening in

the United States could be difficult because screening is recommended in clinical practice guidelines and routinely performed in high-risk patients, randomized trials that compare screening using different imaging modalities or combinations of modalities would be helpful for understanding optimal approaches.

In lieu of such studies, evidence on effects of alternative imaging strategies on intermediate outcomes such as diagnostic thinking, subsequent procedures, and resource utilization could also be informative. Such studies could potentially enroll smaller samples and would probably not require the extended followup needed to assess clinical outcomes.

Although many studies are available on test performance of alternative imaging modalities and strategies, important research gaps remain. Notably, few studies evaluated imaging in true surveillance settings, and evidence on accuracy of imaging for identifying HCC lesions from nonsurveillance settings may not be directly applicable to surveillance due to spectrum effects. More studies are also needed to clarify the role of promising alternative techniques, such as US with contrast, MRI with hepatic-specific contrast, and PET with alternative tracers, on estimates of accuracy. Research should focus on improving methods for identifying small or well-differentiated HCC lesions, for which imaging remains suboptimal.

Conclusions

Based on estimates of test performance, several imaging modalities appear to be reasonable options for surveillance, diagnosis, or staging of HCC. Although there are some potential differences in test performance between different imaging modalities and techniques, more research is needed to understand the effects of such differences on diagnostic thinking and clinical outcomes.

Table A. Summary of evidence on imaging techniques for the surveillance, diagnosis, and staging of hepatocellular carcinoma

Key Question 1. Surveillance
Key Question 1a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Surveillance settings *Unit of analysis: patients with HCC*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.82 (95% CI 0.66 to 0.092, 3 studies) and specificity 0.87 (95% CI 0.77 to 0.93, 2 studies), for a LR+ of 6.2 (95% CI 3.6 to 11) and LR- of 0.20 (0.10 to 0.40).
	CT	Sensitivity: Low Specificity: Low	Sensitivity was 0.84 (95% CI 0.59 to 0.95, 2 studies) and specificity 0.99 (95% CI 0.86 to 0.999, 2 studies).
	MRI or PET	Insufficient	No evidence
Surveillance settings *Unit of analysis: HCC lesions*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.60 (95% CI 0.36 to 0l.80, 2 studies) and specificity 0.94 (95% CI 0.83 to 0.98, 1 study), for a LR+ of 9.8 (95% CI 3.7 to 26) and LR- of 0.43 (95% CI 0.24 to 0.74.
	CT	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.62 (95% CI 0.46 to 0.76, 1 study).
	MRI or PET	No evidence	No evidence
Nonsurveillance settings *Unit of analysis: patients with HCC*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.73 (95% CI 0.46 to 0.90, 8 studies) and specificity 0.93 (95% CI 0.85 to 0.97, 6 studies), for a LR+ of 11 (95% CI 5.4 to 21) and LR- of 0.29 (95% CI 0.13 to 0.65).
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.83 (95% CI 0.75 to 0.89, 16 studies) and specificity 0.92 (95% CI 0.86 to 0.96, 11 studies),

	Imaging Modality or Comparison	Strength of Evidence	Summary
			for a LR+ of 11 (95% CI 5.6 to 20) and LR- of 0.19 (95% CI 0.12 to 0.28).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.87 (95% CI 0.77 to 0.93, 11 studies) and specificity 0.88 (95% CI 0.79 to 0.93, 9 studies), for a LR+ of 7.2 (95% CI 3.9 to 13) and LR- of 0.15 (95% CI 0.08 to 0.27).
	PET	Sensitivity: Low Specificity: Low	FDG PET: Sensitivity was 0.52 (95% CI 0.39 to 0.66, 15 studies) and specificity 0.95 (95% CI 0.92 to 0.99, 5 studies), for a LR+ of 11 (95% CI 2.6 to 49) and LR- of 0.50 (95% CI 0.37 to 0.68). ^{11}C-acetate PET: Sensitivity was 0.85 (95% CI 0.67 to 0.94, 4 studies). Specificity was not reported.
Nonsurveillance settings *Unit of analysis: HCC lesions*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.60 (95% CI 0.42 to 0.75, 11 studies). Only 2 studies reported specificity, with inconsistent results (0.63, 95% CI 0.53 to 0.73 and 0.95, 95% CI 0.85 to 0.99).
	US with contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.76 (95% CI 0.53 to 0.90, 6 studies). No study evaluated specificity.
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.77 (95% CI 0.73 to 0.81, 75 studies) and specificity 0.89 (95% CI 0.83 to 0.93, 20 studies), for a LR+ of 7.0 (95% CI 4.6 to 11) and LR- of 0.25 (95% CI 0.21 to 0.30).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.83 (95% CI 0.80 to 0.86, 69 studies) and specificity 0.83 (95% CI 0.70 to 0.92, 13 studies), for a LR+ of 5.0 and LR- of 0.20 (95% CI 0.16 to 0.2.6).
	PET	Sensitivity: Low Specificity: Low	FDG PET: Sensitivity was 0.56 (95% CI 0.41 to 0.69, 4 studies) and specificity 0.91 (95% CI 0.76 to 0.98, 1 study). ^{11}C-acetate PET: Sensitivity was 0.78 (95% CI 0.61 to 0.89, 4 studies). Specificity was not reported.
Direct (within-study) comparisons of imaging Modalities *Unit of analysis: Patients with HCC*	US without contrast vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.68 (95% CI 0.54 to 0.80) vs. 0.80 (95% CI 0.68 to 0.88), for a difference of -0.12 (95% CI -0.20 to -0.03), based on 6 studies.
	US without contrast vs. MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.61 (95% CI 0.48 to 0.74) vs. 0.81 (95% CI 0.69 to 0.89), for a difference of -0.19 (95% CI -0.30 to -0.08), based on 3 studies.
	MRI vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.88 (95% CI 0.53 to 0.98) vs. 0.82 (95% CI 0.41 to 0.97), for a difference of 0.06 (95% CI -0.05 to 0.17), based on 4 studies.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: HCC lesions*	US without contrast vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.55 (95% CI 0.43 to 0.66) vs. 0.66 (95% CI 0.54 to 0.76) for a difference of -0.11 (95% CI -0.18 to -0.04), based on 3 studies.
	US without contrast vs. MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.57 (95% CI 0.42 to 0.71) vs. 0.79 (95% CI 0.67 to 0.88), for a difference of -0.22 (95% -0.31 to 0.14), based on 3 studies.
	US with contrast vs. CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.58 (95% CI 0.37 to 0.77 vs. 0.74 (95% CI 0.54 to 0.87), for a difference of -0.16 (95% CI -0.32 to 0.01), based on 3 studies.
	US with contrast vs. MRI	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.25 to 0.80) vs. 0.70 (95% CI 0.40 to 0.89), for a difference of -0.16 (95% CI -0.30 to -0.02), based on 2 studies.
	MRI vs. CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.81 (95% CI 0.77 to 0.84) vs. 0.72 (95% CI 0.67 to 0.77), for a difference of 0.09 (95% CI 0.06 to 0.12), based on 28 studies. Findings were similar when studies were stratified according to use of non-hepatic specific or hepatic specific contrast.

	Imaging Modality or Comparison	Strength of Evidence	Summary
Multiple imaging modalities	Various combinations	Sensitivity: Low	1 study found sensitivity of imaging with various combinations of two imaging modalities was similar or lower than single modality imaging, based on concordant positive findings on 2 imaging modalities. The other study reported higher sensitivity with multiple imaging modalities than with single modality imaging, but criteria for positive results based on multiple imaging modalities were unclear

Key Question 1a.i. Effects of reference standard on test performance (based on HCC lesions as the unit of analysis)

Imaging Modality or Comparison	Strength of Evidence	Summary
US	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.34 (95% CI 0.21 to 0.49) in 5 studies that used explanted liver as the reference standard and ranged from 0.70 to 0.85 in studies that used other reference standards.
CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.69 (95% CI 0.60 to 0.77) in 21 studies that used explanted liver as the reference standard and ranged from 0.79 to 0.85 in studies that used other reference standards.
MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.69 (95% CI 0.59 to 0.77) in 15 studies that used explanted liver as the reference standard and ranged from 0.85 to 0.88 in studies that used other reference standards.
PET	Sensitivity: Low Specificity: Insufficient	No study of FDG PET used an explanted liver reference standard.

Key Question 1a.ii. Effects of patient, tumor, technical, and other factors on test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Lesion size	US	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.78 to 0.94) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions <2 cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies. The difference was larger in studies of US without contrast than studies of US with contrast, but these findings are difficult to interpret because sensitivity for HCC lesions <20 mm was much lower in the studies of US without contrast. For US without contrast, sensitivity was 0.09 (95% CI 0.02 to 0.29, 4 studies) for lesions < 10 mm to 0.50 (95% CI 0.23 to 0.78, 4 studies) for lesions 10 to 20 mm and 0.88 (95% CI 0.66 to 0.96, 4 studies) for lesions >20 mm , for a difference of 0.37 (95% CI 0.18 to 0.57) for lesions >20 mm vs. 10 to 20 mm, and 0.41 (95% CI 0.19 to 0.63) for lesions 10 to 20 mm vs. <10 mm. For ultrasound with contrast, three studies found sensitivity of 0.64 (95% CI 0.33 to 0.87) for lesions 10 to 20 mm and 0.91 (95% CI 0.71 to 0.98) for lesions >20 mm, for a difference of 0.26 (95% CI 0.04 to 0.48).
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.91 to 0.95) for lesions >2 cm and 0.62 (95% CI 0.56 to 0.68) for lesions <2 cm, for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36), based on 33 studies. Sensitivity was 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10 to 20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 vs. 10 to 20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10 to 20 vs. <10 mm.

	Imaging Modality or Comparison	Strength of Evidence	Summary
	MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.96 (95% CI 0.93 to 0.97) for lesions >2 cm and 0.65 (95% CI 0.57 to 0.73) for lesions <2 cm, for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37), based on 25 studies. Sensitivity was 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10 to 20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 vs. 10 to 20 mm and 0.34 (95% CI 0.27 to 0.41) for 10 to 20 vs. <10 mm.
	PET	Sensitivity: Low	For FDG PET, sensitivity was consistently higher for larger lesions, based on 5 studies. Data were not pooled due to differences in the tumor size categories evaluated. Two studies of ^{11}C-acetatate PET found inconsistent effects of lesion size on sensitivity
Degree of tumor differentiation	US with contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.83 (95% CI 0.55 to 0.95) for moderately or poorly-differentiated HCC lesions and 0.43 (95% CI 0.15 to 0.76) for well differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.17 to 0.64), based on 3 studies.
	CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.82 (95% CI 0.66 to 0.91) for moderately or poorly-differentiated HCC lesions and 0.50 (95% CI 0.29 to 0.70) for well differentiated lesions, for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45), based on 5 studies.
	MRI	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.26 to 0.79) for moderately or poorly-differentiated HCC lesions and 0.38 (95% CI 0.17 to 0.64) for well differentiated lesions, but the difference was not statistically significant (0.16, 95% CI -0.11 to 0.43), based on 2 studies.
	PET	Sensitivity: Low Specificity: Insufficient	For FDG PET, sensitivity was consistently higher for more poorly-differentiated lesions than more well differentiated lesions, based on five studies. In three studies of ^{11}C-acetate PET and one study of ^{18}F-fluorochlorine, sensitivity for more well differentiated lesions was not lower than more poorly-differentiated lesions.
Other factors	US	Low	In 2 studies that directly compared US with versus without contrast, there was no clear difference in sensitivity (-0.04, 95% CI -0.11 to 0.04). 1 study that directly compared use of Doppler versus no Doppler showed no clear effect on estimates of sensitivity. Lesion depth and body mass index had no effect on estimates of sensitivity.
	CT	Low-Moderate	Using patients with HCC as the unit of analysis, studies with a contrast rate ≥3 ml/s reported a higher sensitivity (0.87, 95% CI 0.77 to 0.93, 8 studies) than studies with a contrast rate <3 ml/s (0.71, 95% CI 0.50 to 0.85, 4 studies) and studies with delayed phase imaging reported somewhat higher sensitivity (0.89, 95% CI 0.81 to 0.94, 7 studies) than studies without delayed phase imaging (0.74, 95% CI 0.66 to 0.87, 7 studies), but there were no clear effects in studies that used HCC lesions as the unit of analysis.
	MRI	Low-Moderate	There were no clear differences in estimates of diagnostic accuracy when studies were stratified according to MRI scanner type (1.5 vs. 3.0 T), imaging phases evaluated (with or without delayed phase imaging), timing of delayed phase imaging (>120 seconds vs. <120 seconds), section thickness (≤5 mm for enhanced images vs. >5 mm), or use of diffusion-weighted imaging. In studies that directly compared diagnostic accuracy with different types of contrast, hepatic-specific contrast agents were associated with slightly higher sensitivity than non-hepatic specific contrast agents (0.82, 95% CI 0.71

	Imaging Modality or Comparison	Strength of Evidence	Summary
			to 0.90 vs. 0.75, 95% CI 0.61 to 0.85, difference 0.07, 95% CI 0.01 to 0.14, 5 studies).
	PET	Low-Moderate	FDG PET was associated with lower sensitivity that ^{11}C-acetate PET when either patients (0.58 vs. 0.81, for a difference of -0.23, 95% CI -0.34 to -0.13, 3 studies**)** or HCC lesions (0.52 vs. 0.79, for a difference of -0.27, 95 %CI -0.36 to -0.17, 3 studies) were the unit of analysis. FDG PET was also associated with lower sensitivity that dual tracer PET with FDG and ^{11}C-acetate or 18F-choline PET, but evidence was limited to 1 or 2 studies for each of these comparisons. Using patients as the unit of analysis, sensitivity of FDG PET (0.39, 95% CI 0.24 to 0.56, 8 studies) was lower than sensitivity of FDG PET/CT (0.65, 95% CI 0.50 to 0.78, 7 studies).

Key Question 1b. Diagnostic thinking

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 1c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
US plus serum AFP	Low	1 cluster randomized controlled trial (n=18816) conducted in China found screening every 6 months with noncontrast US plus serum AFP versus no screening in persons 35 to 79 years of age (mean 42 years) with HBV infection or chronic hepatitis without HBV infection associated with lower risk of HCC-related mortality (32 vs. 54 deaths, rate ratio 0.63, 95% CI 0.41 to 0.98) at 5 year followup, but was rated high risk of bias. 2 trials found no clear differences in mortality with US screening at 4- vs. 12-month intervals, or at 3- vs. 6-month intervals.

Key Question 1d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
MRI, CT, US	Insufficient	1 study reported no serious adverse events associated with administration of gadoxetic acid for MRI and one study reported no clear differences in adverse events between CT with contrast at 3 ml/s versus 5 ml/s. No study reported rates of adverse events associated with use of microbubble contrast agents in US, and harms were not reported in randomized trials of screening with imaging.

Key Question 2. Diagnosis

Key Question 2a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Evaluation of a previously identified lesion *Unit of analysis: Patients with HCC*	US with contrast	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.79 to 0.94, 8 studies) and specificity 0.92 (95% CI 0.84 to 0.96, 5 studies), for a LR+ of 11 (95% CI 5.5 to 20) and LR- of 0.13 (95% CI 0.07 to 0.24).
	US without contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.78 (95% CI 0.72 to 0.83) in 2 studies; specificity was not reported.
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.85 (95% CI 0.76 to 0.91, 5 studies) and specificity 0.92 (95% CI 0.86 to 0.96, 3 studies), for a LR+ of 11 (95% CI 5.7 to 22) and LR- of 0.17 (95% CI 0.10 to 0.27).

	Imaging Modality or Comparison	Strength of Evidence	Summary
	MRI	Sensitivity: Low Specificity: Low	Sensitivity was 0.76 (95% CI 0.62 to 0.86, 3 studies) and specificity 0.87 (95% CI 0.70 to 0.95, 3 studies), for a LR+ of 5.9 (95% CI 2.5 to 14) and LR- of 0.28 (95% CI 0.18 to 0.43).
Evaluation of a previously identified lesion *Unit of analysis: HCC lesions*	US with contrast	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.86 (95% CI 0.79 to 0.91, 21 studies) and specificity 0.93 (95% CI 0.87 to 0.96, 11 studies) for a LR+ of 12 (95% CI 6.3 to 21) and LR- of 0.15 (95% CI 0.10 to 0.23).
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.80 (95% CI 0.67 to 0.88, 12 studies) and specificity 0.89 (95% CI 0.29 to 0.99, 6 studies), for a LR+ of 6.9 (95% CI 0.53 to 91) and LR- of 0.23 (95% CI 0.13 to 0.40).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.79 (95% CI 0.69 to 0.87, 13 studies) and specificity 0.95 (95% CI 0.82 to 0.99, 12 studies), for a LR+ of 15 (95% CI 4.4 to 50) and LR- of 0.22 (95% CI 0.15 to 0.33).
	PET	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.56 to 0.57 and specificity 1.0 in 2 studies of FDG PET.
For distinguishing HCC lesions from non-HCC hepatic lesions	US with contrast	Low	1 study found US with sulfur hexafluoride contrast associated with a sensitivity of 0.94 (62/66) and specificity of 0.68 (23/34) for distinguishing hypervascular HCC from focal nodular hyperplasia, using quantitative methods.
	CT	Low	4 studies evaluated accuracy of CT for distinguishing HCC from non-HCC lesions, but the non-HCC lesions varied in the studies, precluding strong conclusions.
	MRI	Moderate	4 studies reported inconsistent results for distinguishing small (<2 to 3 cm) hypervascular HCC lesions from hypervascular pseudolesions, with sensitivity 0.47 and 0.52 in 2 studies, and 0.91 and 0.92 in the other two. Specificity was 0.93 or higher in all four studies. Five other studies evaluated accuracy of MRI for distinguishing HCC from other non-HCC lesions, but the non-HCC lesions varied in the studies.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: Patients with HCC*	US without contrast vs. CT	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.78 (95% CI 0.70 to 0.85) vs. 0.89 (95% CI 0.84 to 0.95), for a difference of -0.12 (95% CI -0.21 to -0.02), based on 1 study.
	US with contrast vs. CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.91 (0.85 to 0.95) vs. 0.87 (95% CI 0.79 to 0.92), for a difference of 0.04 (95% CI -0.01 to 0.10), based on 4 studies.
	MRI vs. CT	Sensitivity: Low Specificity: Low	Sensitivity was 0.81 (95% CI 0.70 to 0.92) vs. 0.74 (95% CI 0.62 to 0.87), for a difference of 0.06 (-0.10 to 0.23), based on 1 study.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: HCC lesion*	US with contrast vs. CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.89 to 0.97) vs. 0.91 (95% CI 0.85 to 0.94), for a difference of 0.03 (95% CI -0.03 to 0.09), based on 3 studies.
	US with contrast vs. MRI	Sensitivity: Low Specificity: Low	Sensitivity was 0.79 (95% CI 0.65 to 0.94) vs. 0.83 (95% CI 0.69 to 0.97), for a difference of -0.03 (95% CI -0.24 to 0.17), based on 1 study.
	MRI vs. CT	Sensitivity: Low Specificity: Low	One study found MRI associated with higher sensitivity (0.84, 95% CI 0.76 to 0.92 vs. 0.62, 95% CI 0.52 to 0.72, for a difference of 0.22, 95% CI 0.09 to 0.35) but lower specificity (0.36, 95% CI 0.20 to 0.52 vs. 0.72, 95% CI 0.58 to 0.87, for a difference of -0.36, 95% CI -0.58 to -0.15) than CT.
Multiple imaging modalities	Various combinations	Moderate	In 4 studies in which positive results with multiple modality imaging were defined as concordant typical findings for HCC on 2 imaging modalities, sensitivity was lower than with a single modality (difference in sensitivity ranged from 0.09 to 0.27), with no clear difference in

	Imaging Modality or Comparison	Strength of Evidence	Summary
			specificity. In three studies in which positive results with multiple modality imaging were defined as typical findings for HCC on at least one of the imaging techniques, sensitivity was higher than with a single modality (increase in sensitivity ranged from 0.09 to 0.25), with no clear difference in specificity. 1 study found that a sequential imaging strategy in which a second imaging test was only performed for indeterminant results on initial CT increased sensitivity for HCC from 0.53 to 0.74 to 0.79.

Key Question 2a.i. Effects of reference standard on test performance (based on HCC lesions as the unit of analysis)

Imaging Modality or Comparison	Strength of Evidence	Summary
All	Sensitivity: Moderate Specificity: Moderate	No study used explanted liver as the reference standard. There were no clear differences across imaging modalities in estimates of diagnostic accuracy in analyses stratified by use of different non-explant reference standards.

Key Question 2a.ii. Effects of patient, tumor, technical, and other factors on test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Lesion size	US	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.78 to 0.94) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions > 2cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies.
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.91 to 0.95) for lesions >2 cm and 0.62 (95% CI 0.56 to 0.68) for lesions <2 cm, for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36), based on 33 studies. Sensitivity was 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10 to 20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 vs. 10 to 20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10 to 20 vs. <10 mm.
	MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.96 (95% CI 0.93 to 0.97) for lesions >2 cm and 0.65 (95% CI 0.57 to 0.73) for lesions <2 cm, for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37), based on 25 studies. Sensitivity was 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10 to 20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 vs. 10 to 20 mm and 0.34 (95% CI 0.27 to 0.41) for 10 to 20 vs. <10 mm.
Degree of tumor differentiation	US	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.84 (95% CI 0.64 to 0.94) for moderately or poorly-differentiated HCC lesions and 0.43 (95% CI 0.21 to 0.69) for well differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.22 to 0.59), based on 4 studies.
	CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.82 (95% CI 0.66 to 0.91) for moderately or poorly-differentiated HCC lesions and 0.50 (95% CI 0.29 to 0.70) for well differentiated lesions, for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45), based on 5 studies.

	Imaging Modality or Comparison	Strength of Evidence	Summary
	MRI	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.26 to 0.79) for moderately or poorly-differentiated HCC lesions and 0.38 (95% CI 0.17 to 0.64) for well differentiated lesions, but the difference was not statistically significant (0.16, 95% CI -0.11 to 0.43), based on 2 studies.
Other factors	US	Insufficient-Low	In 2 studies that directly compared US with versus without contrast, US with contrast was associated with sensitivity of 0.89 (95% CI 0.83 to 0.93) and US without contrast with a sensitivity of (0.39) 95% CI 0.32 to 0.47), for a difference in sensitivity of 0.50 (95% CI 0.41 to 0.58). Based on across-study comparisons, there were no clear differences in sensitivity between different US contrast agents; no study directly compared different contrast agents. There were no differences in sensitivity of US based on lesion depth (3 studies) or body mass index (2 studies).
	CT	Insufficient-Low	Evidence on effects of technical parameters (type of CT scanner, use of delayed phase imaging, section thickness) was limited by small numbers of studies with wide confidence intervals and methodological limitations, precluding reliable conclusions. 2 studies found no clear difference in sensitivity of CT for HCC in patients with versus without cirrhosis.
	MRI	Low-Moderate	There were no clear differences in estimates of sensitivity based on the type of MRI machine (3.0 T vs. 1.5 T), type of contrast, use of delayed phase imaging, timing of delayed phase imaging, and section thickness. Estimates were similar were studies that used diffusion-weighted imaging were excluded.

Key Question 2b. Diagnostic thinking

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 2c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 2d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
US and CT	Insufficient	1 study of US (with and without contrast) and CT reported harms, but did not stratify results by imaging technique. The overall rate of adverse drug-related events was 10%, with all events classified as mild.

Key Question 3. Staging

Key Question 3a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Staging accuracy, using TNM criteria	CT	Moderate	The proportion correctly staged ranged from 28% to 58%, the proportion overstaged from 2% to 27%, and the proportion understaged from 25% to 52%, based on 6 studies.
	MRI	Low	The proportion correctly staged were 10% and 31%, the proportion overstaged 10% and 31%, and the proportion understaged 29% and 31%, based on 2

			studies.
	PET	Low	1 study found 26% of patients were correctly staged with FDG PET and 91% with ^{11}C-choline PET.
	MRI vs. CT	Low	2 studies reported similar staging accuracy.
Identification of metastatic disease *Unit of analysis: Patients with metastatic HCC*	PET	Sensitivity: Moderate Specificity: Moderate	Sensitivity of FDG PET was 0.85 (95% CI 0.71 to 0.93, 6 studies) and specificity 0.93 (95% CI 0.89 to 0.95, 5 studies), for a LR+ of 11 (95% CI 7.8 to17) and LR- of 0.16 (95% CI 0.08 to 0.33). 1 study that directly compared sensitivity of FDG PET to 11-chloroacetate PET reported comparable sensitivity (0.79 vs. 0.71), though sensitivity was higher when both tracers were used (0.98).
Identification of metastatic disease *Unit of analysis: Metastatic HCC lesions*	PET	Sensitivity: Moderate Specificity: Insufficient	Sensitivity of FDG PET was 0.82 (95% CI 0.72 to 0.90, 5 studies). 1 study that directly compared sensitivity of FDG to 11-choloroacetate PET reported comparable sensitivity (0.86 vs. 0.77, respectively).

Key Question 3.a.i. Effects of reference standard on test performance

Imaging Modality or Comparison	Strength of Evidence	Summary
CT, MRI, PET	Sensitivity: Low Specificity: Low	Evidence was insufficient to determine effect s of different reference standards on accuracy of staging using TNM criteria or accuracy of PET for identifying metastatic HCC because few studies evaluated alternative reference standards.

Key Question 3.a.ii. Effects of patient, tumor, and technical factors on test performance

Imaging Modality or Comparison	Strength of Evidence	Summary
CT, MRI, PET	No evidence	For accuracy of staging using TNM criteria, no study evaluated effects of patient-level characteristics or other factors on accuracy of imaging techniques for staging.
PET	Low-Moderate	In 1 study that directly compared sensitivity of PET vs. PET/CT for identifying metastatic HCC lesions, there was no clear difference in sensitivity. 4 studies of FDG PET found sensitivity increased as lesion size increased, but the number of lesions <1 cm was small (total of 20). 8 studies generally found sensitivity of FDG PET higher for lymph and bone metastasis than for lung metastasis, but samples were small, precluding strong conclusions.

Key Question 3b. Diagnostic thinking

	Imaging Modality or Comparison	Strength of Evidence	Summary
Transplant eligibility, using Milan criteria	CT	Moderate	The proportion correctly assessed for transplant eligibility ranged from 40% to 96%. The proportion of patients who met transplant criteria based on CT but exceeded criteria based on the reference standard was 3.5 to 7.8%, based on 3 studies. 2 studies found that 2.3% and 16% of patients who underwent transplantation based on Milan criteria had no HCC lesions on examination of explanted livers.
	CT vs. MRI	Low	1 study reported similar accuracy.
	PET vs. CT	Low	1 study found ^{11}C-choline PET more accurate than CT (95% vs. 40%).

Use of resection and ablative therapies	MRI vs. CT	Low	1 study reported that the proportion of decisions to perform resection or ablative therapies that were classified as correct were similar for MRI (90% and 90%, respectively) and CT (80% and 77%, respectively).

Key Question 3c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
US with contrast vs. US without contrast plus CT	Low	1 cohort study found that contrast enhanced US identified more small (≤2 cm) HCC lesions than noncontrast US plus CT (36 vs. 31), and was associated with a higher complete necrosis rate following ablation (92% or 106/115 vs. 83% or 93/112 lesions, p=0.036), but was rated high risk of bias.

Key Question 3d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

AFP = alpha-fetoprotein; CT = computed tomography; FDG = fludeoxyglucose; HBV = hepatitis B virus; HCC = hepatocellular carcinoma; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; MRI = magnetic resonance imaging; PET = positron emission tomography; TNM = Tumor, Node Metastasis Staging; US = ultrasound

References

1. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol. 2001 Sep;2(9):533-43. PMID: 11905707.
2. National Cancer Institute. Liver Cancer. http://www.cancer.gov/cancertopics/types/liver. Accessed on September 21 2011.
3. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology. 2011 Mar;53(3):1020-2. PMID: 21374666.
4. Cabibbo G, Enea M, Attanasio M, et al. A meta-analysis of survival rates of untreated patients in randomized clinical trials of hepatocellular carcinoma. Hepatology. 2010 Apr;51(4):1274-83. PMID: 20112254.
5. Forner A, Reig ME, de Lope CR, et al. Current strategy for staging and treatment: the BCLC update and future prospects. Semin Liver Dis. 2010 Feb;30(1):61-74. PMID: 20175034.
6. Mazzaferro V, Bhoori S, Sposito C, et al. Milan criteria in liver transplantation for hepatocellular carcinoma: An evidence-based analysis of 15 years of experience. Liver Transpl. 2011;17(S2):S44-S57.
7. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(13)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2013. Chapters available at: www.effectivehealthcare.ahrq.gov.
8. Chang SM, Matchar DB, eds. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2012. Available at http://effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.
9. Submit Scientific Information Packets. Rockville, MD: Agency for Healthcare Research and Quality; 2013. http://effectivehealthcare.ahrq.gov/index.cfm/submit-scientific-information-packets/. Accessed on November 12, 2013.
10. Wald C, Russo MW, Heimbach JK, et al. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology. 2013 Feb;266(2):376-82. PMID: 23362092.
11. ESP Reports in Progress. Washington, DC: U.S. Department of Veterans Affairs. Health Services Research & Development Services.; 2013. http://www.hsrd.research.va.gov/publications/esp/in_progress.cfm. Accessed on November 11, 2013.
12. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001 Apr;20(3 Suppl):21-35. PMID: 11306229.
13. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36. PMID: 22007046.
14. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(13)-EHC063-EF. Chapter 9. Rockville, MD: Agency for Healthcare Research and Quality; 2013. Available at http://www.effectivehealthcare.ahrq.gov/.
15. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013 Apr 2;158(7):544-54. PMID: 23546566.
16. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1198-207. PMID: 21463926.
17. Chalasani N, Horlander JC, Sr., Said A, et al. Screening for hepatocellular carcinoma in patients with advanced cirrhosis. Am J Gastroenterol. 1999 Oct;94(10):2988-93. PMID: 10520857.
18. Van Thiel DH, Yong S, Li SD, et al. The development of de novo hepatocellular carcinoma in patients on a liver transplant list: frequency, size, and assessment of current screening methods. Liver Transpl. 2004 May;10(5):631-7. PMID: 15108254.
19. Goto E, Masuzaki R, Tateishi R, et al. Value of post-vascular phase (Kupffer imaging) by contrast-enhanced ultrasonography using Sonazoid in the detection of hepatocellular carcinoma. J Gastroenterol. 2012;47(4):477-85. PMID: 22200940.

20. Kunishi Y, Numata K, Morimoto M, et al. Efficacy of fusion imaging combining sonography and hepatobiliary phase MRI with Gd-EOB-DTPA to detect small hepatocellular carcinoma. AJR Am J Roentgenol. 2012 Jan;198(1):106-14. PMID: 22194485.

21. Lencioni R, Piscaglia F, Bolondi L. Contrast-enhanced ultrasound in the diagnosis of hepatocellular carcinoma. J Hepatol. 2008 May;48(5):848-57. PMID: 18328590.

22. Naaktgeboren CA, de Groot JA, van Smeden M, et al. Evaluating diagnostic accuracy in the face of multiple reference standards. Ann Intern Med. 2013 Aug 6;159(3):195-202. PMID: 23922065.

23. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004 Jul;130(7):417-22. PMID: 15042359.

24. Colli A, Fraquelli M, Casazza G, et al. Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. Am J Gastroenterol. 2006 Mar;101(3):513-23. PMID: 16542288.

25. Westwood M, Joore M, Grutters J, et al. Contrast-enhanced ultrasound using SonoVue(R) (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2013 Apr;17(16):1-243. PMID: 23611316.

26. Guang Y, Xie L, Ding H, et al. Diagnosis value of focal liver lesions with SonoVue-enhanced ultrasound compared with contrast-enhanced computed tomography and contrast-enhanced MRI: a meta-analysis. J Cancer Res Clin Oncol. 2011 Nov;137(11):1595-605. PMID: 21850382.

27. Xie L, Guang Y, Ding H, et al. Diagnostic value of contrast-enhanced ultrasound, computed tomography and magnetic resonance imaging for focal liver lesions: a meta-analysis. Ultrasound Med Biol. 2011 Jun;37(6):854-61. PMID: 21531500.

28. Lin C-Y, Chen J-H, Liang J-A, et al. 18F-FDG PET or PET/CT for detecting extrahepatic metastases or recurrent hepatocellular carcinoma: a systematic review and meta-analysis. Eur J Radiol. 2012 Sep;81(9):2417-22. PMID: 21899970.

29. Aghoram R, Cai P, Dickinson JA. Alpha-foetoprotein and/or liver ultrasonography for screening of hepatocellular carcinoma in patients with chronic hepatitis B. Cochrane Database Syst Rev. 2012;9.

30. Lachs MS, Nachamkin I, Edelstein PH, et al. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. Ann Intern Med. 1992 Jul 15;117(2):135-40. PMID: 1605428.

Background and Objectives

Hepatocellular carcinoma (HCC) is the most common primary malignant neoplasm of the liver, usually developing in individuals with chronic liver disease or cirrhosis. Worldwide, it is the fifth most common cause of cancer and the third most common cause of cancer death.[1] According to the National Cancer Institute, there were 156,940 deaths attributed to liver and intrahepatic bile duct cancer in the United States in 2011, with 221,130 new cases diagnosed.[2] The National Cancer Institute's Surveillance, Epidemiology, and End Results Cancer Statistics Review found that the lifetime risk of developing liver and intrahepatic bile duct cancer is 1 in 132, with the age-adjusted incidence rate being 7.3 per 100,000 people per year.[3] The highest incidence rates in the United States are found in Asian/Pacific Islanders (22.1 per 100,000 men and 8.4 per 100,000 women). The age-adjusted death rate is estimated at 5.2 per 100,000 people per year in the United States, with the highest sex-specific rates among Asian/Pacific Islander men (14.7 per 100,000) and American Indian/Alaskan Native women (6.6 per 100,000). The overall 5-year relative survival rate is 14.4 percent.

The 2011 Annual Report to the Nation on the Status of Cancer reported that deaths from liver cancer significantly increased from 1998–2007 in both men and women.[4] The increase was mostly attributable to cirrhosis due to hepatitis B virus (HBV) infection, hepatitis C virus (HCV) infection, or long-term alcohol abuse, with HCV infection accounting for at least half of the observed increase[5,6] In the United States, HCV infection is the most frequently identified cause of HCC, which is present in about half of all cases, though 15 to 50 percent of patients have no identifiable etiology.[6] Worldwide, HBV infection is responsible for the majority of HCC cases, particularly in developing countries,[7] though the incidence of HBV infection and associated complications has declined following to the widespread implementation of universal vaccination programs.[8] The American Association for the Study of Liver Diseases (AASLD) recommends surveillance for the following groups at high risk for developing HCC: Asian male HBV carriers over age 40, Asian female HBV carriers over age 50, HBV carriers with a family history of HCC, African/North American black HBV carriers, HBV or HCV carriers with cirrhosis, all individuals with other causes for cirrhosis (including alcoholic cirrhosis), and patients with stage 4 primary biliary cirrhosis.[9]

HCC is an aggressive tumor associated with poor survival without treatment.[10] However, when diagnosed early, HCC may be amenable to potentially curative therapy. The three phases of pretherapy evaluation of HCC include surveillance, diagnosis, and staging.[9] Surveillance is the use of periodic testing to identify lesions in the liver that are clinically suspicious for HCC.[9] The diagnosis phase involves the use of additional tests (radiological and/or histopathological) to confirm that the lesion detected in the liver is indeed HCC. Once the lesion is confirmed as HCC, staging is important for informing prognosis and treatment decisions. A number of staging systems are available, including the traditional TNM (tumor, node, metastasis) classification, based on the size, number, and location of primary lesions, the presence of invasion into vascular and biliary structures, and the presence of regional nodal and distant metastases.[11] More recently, the Barcelona Clinic Liver Cancer (BCLC) staging system,[12] which incorporates additional factors associated with prognosis such as liver functional status, physical status, cancer related symptoms, and impact of treatment, has become the de facto staging reference standard.[9] To select patients who are suitable for liver transplantation, the Milan criteria (one lesion <5 cm or up to 3 lesions <3 cm, with no extrahepatic manifestations or vascular invasion) are used to identify patients likely to experience better posttransplantation outcomes, though other methods have been proposed.[13]

A number of imaging techniques are available to identify the presence of lesions, diagnose HCC, and determine the stage of the disease (Table 1). These include ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). The typical use of each of these imaging modalities varies. For example, PET scan is typically used for staging and identification of metastatic disease, but not for surveillance and diagnosis. US without contrast is the most frequently used modality for surveillance and recommended by the AASLD for this purpose.[9] Because HCC is typically a hypervascular lesion, arterial enhancing contrasts are frequently used to increase the sensitivity and specificity of imaging techniques such as CT or MRI. Similarly, microbubble-enhanced US is performed to evaluate liver lesions in regions of the world such as Europe and Asia, though agents are not yet approved by the U.S. Food and Drug Administration (FDA) for this purpose.[14] Because the microbubbles are present for only a limited period of time in the liver, such that a comprehensive evaluation of the entire liver is not possible, contrast-enhanced US is typically performed for the targeted evaluation and characterization of focal liver lesions previously identified on US without contrast or other imaging studies rather than for surveillance of the entire liver.

Understanding the diagnostic accuracy of imaging methods and how they affect clinical decisionmaking and, ultimately, patient outcomes is a challenge. Imaging techniques may be used alone, in various combinations or algorithms, and/or with liver-specific biomarkers, resulting in many potential comparisons. In addition, surveillance and diagnostic strategies vary. For example, some centers use periodic US alone for surveillance, while others use US alternatively with either CT or MRI every 6 months, with or without use of biomarkers such as alpha-fetoprotein. Technical aspects of imaging methods are complex and continuously evolving. Published standards for CT and MRI are available, providing guidance regarding minimum recommended technical specifications with regard to scanner types, section thickness, imaging phases, timing of imaging phases, and other factors (Tables 2 and 3).[15] Other technical variations have also been introduced, including MRI with the liver-specific contrast agents[16] such as gadobenate or gadoxetic acid disodium (rather than standard nonspecific contrast agents such as gadodiamide or gadopentetate),[17] CT utilizing dual energy source or spectral techniques,[18,19] US with use of contrast enhancement,[20] and PET with use of tracers such as ^{18}F-fluorothymidine (FLT) or ^{11}C-choline,[21] rather than the standard ^{18}F-fluorodeoxyglucose (FDG). The use of different reference standards—such as explanted liver specimens from patients undergoing transplantation, percutaneous or surgical biopsy, imaging, clinical followup, or combinations of these methods—could introduce heterogeneity. Reference standards also are susceptible to misclassification due to sampling error, inadequate specimens, insufficient followup, or other factors. Finally, other considerations, including risk factors for HCC and lesion characteristics, such as tumor size or degree of differentiation, severity of hepatic fibrosis, and etiology of liver disease, may impact the diagnostic accuracy or clinical utility of imaging strategies.

In addition to imaging studies, serological biomarkers for HCC can be used to aid in diagnosis. Alpha-fetoprotein is the most widely used serological marker for HCC surveillance, but recommended only as an adjunct to imaging due to limited sensitivity and specificity.[9] A newer biomarker is des-gamma-carboxy prothrombin, though its role in the surveillance and early diagnosis of HCC has not yet been defined.[22] Other biomarkers, such as glypican 3, heat shock protein 70, and glutamine synthetase, have not been validated in the clinical setting and are not currently recommended for use in screening.[9,23]

Accurate identification and staging of HCC is critical for providing optimal patient care. However, clinical uncertainty remains regarding optimal imaging strategies, due to the factors

described above. The purpose of this report is to comprehensively review the comparative effectiveness and diagnostic performance of different imaging modalities and strategies for surveillance, diagnosis, and staging of HCC.

Scope and Key Questions

The Key Questions and corresponding analytic frameworks used to guide this report are shown below. Separate analytic frameworks address surveillance (Key Question 1, Figure 1), diagnosis (Key Question 2, Figure 2), and staging (Key Question 3, Figure 3). The analytic frameworks show the target populations, interventions (imaging tests), and outcomes (diagnostic accuracy, diagnostic thinking, clinical outcomes, and harms) that we examined.

Key Question 1. What is the comparative effectiveness of available imaging-based surveillance strategies, used singly or in sequence, for detecting hepatocellular carcinoma (HCC) among individuals undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for HCC)?

e. What is the comparative test performance of imaging-based surveillance strategies for detecting HCC?
    i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
    ii. How is the comparative effectiveness modified by patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, or other factors (e.g., results of biomarker tests, setting)?
f. What is the comparative effectiveness of imaging-based surveillance strategies on intermediate outcomes like diagnostic thinking?
g. What is the comparative effectiveness of imaging-based surveillance strategies on clinical and patient-centered outcomes?
h. What are the adverse effects or harms associated with imaging-based surveillance strategies?

Key Question 2. What is the comparative effectiveness of imaging techniques, used singly, in combination, or in sequence, in diagnosing HCC among individuals in whom an abnormal lesion has been detected while undergoing surveillance for HCC (individuals at high risk for HCC and individuals who have undergone liver transplants for HCC) or through the evolution of symptoms and abdominal imaging done for other indications?

e. What is the comparative test performance of imaging techniques for diagnosing HCC?
    i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical imaging and followup?
    ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?
f. What is the comparative effectiveness of the various imaging techniques on intermediate outcomes such as diagnostic thinking and use of additional diagnostic procedures such as fine-needle or core biopsy?
g. What is the comparative effectiveness of the various imaging techniques on clinical and patient-centered outcomes?
h. What are the adverse effects or harms associated with imaging-based diagnostic strategies?

Key Question 3. What is the comparative effectiveness of imaging techniques, used singly, in combination, or in sequence in staging HCC among patients diagnosed with HCC?

e. What is the comparative test performance of imaging techniques to predict HCC tumor stage?
    i. How is a particular technique's test performance modified by use of various reference standards (e.g., explanted liver samples, histological diagnosis, or clinical and imaging followup)?
    ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?
f. What is the comparative test performance of imaging techniques on diagnostic thinking?
g. What is the comparative effectiveness of imaging techniques on clinical and patient-centered outcomes?
h. What are the adverse effects or harms associated with imaging-based staging strategies?

Figure 1. Analytic framework—surveillance (Key Question 1)

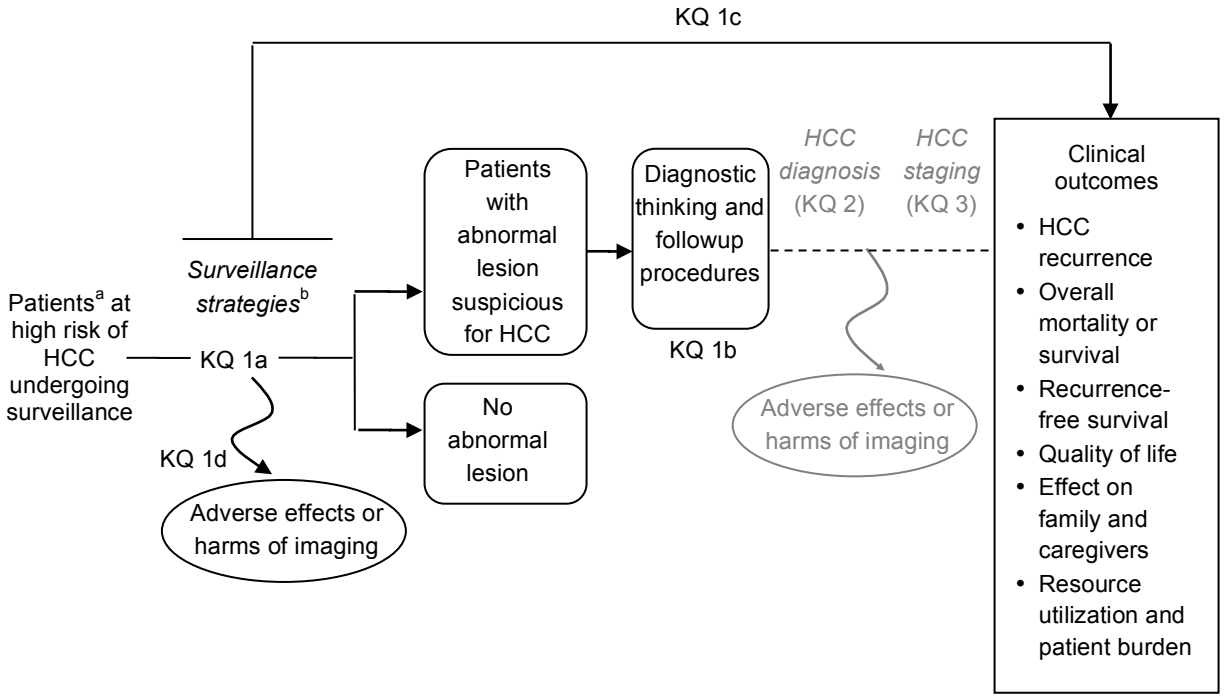
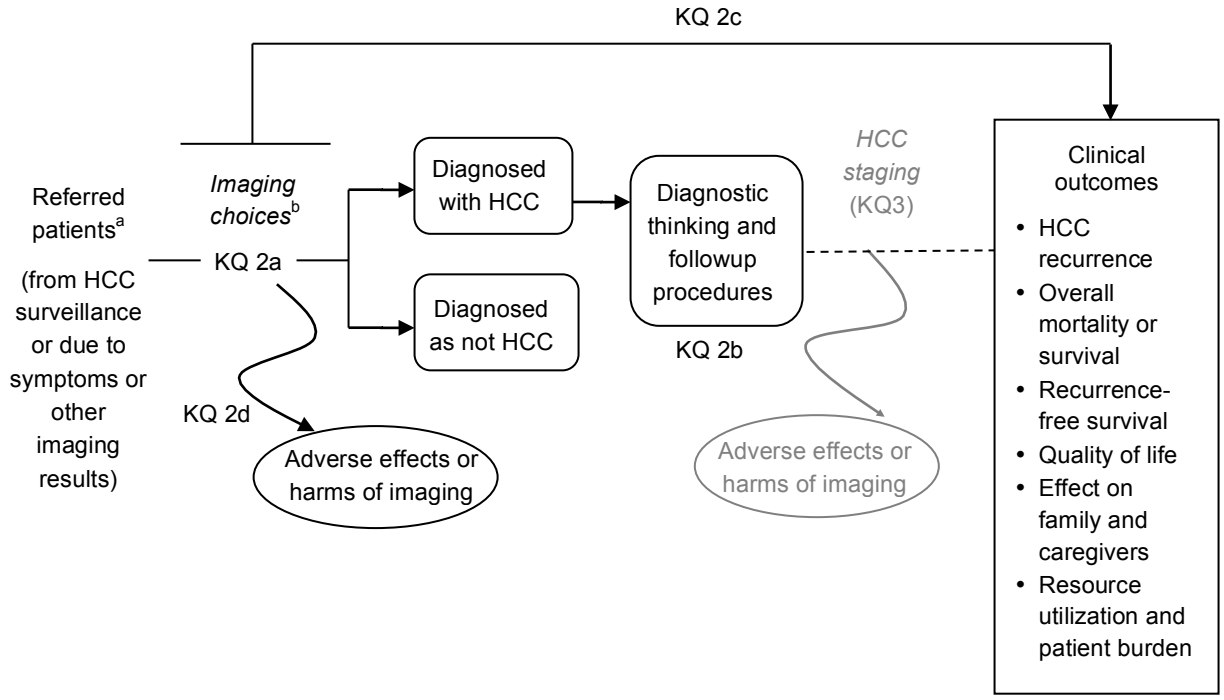
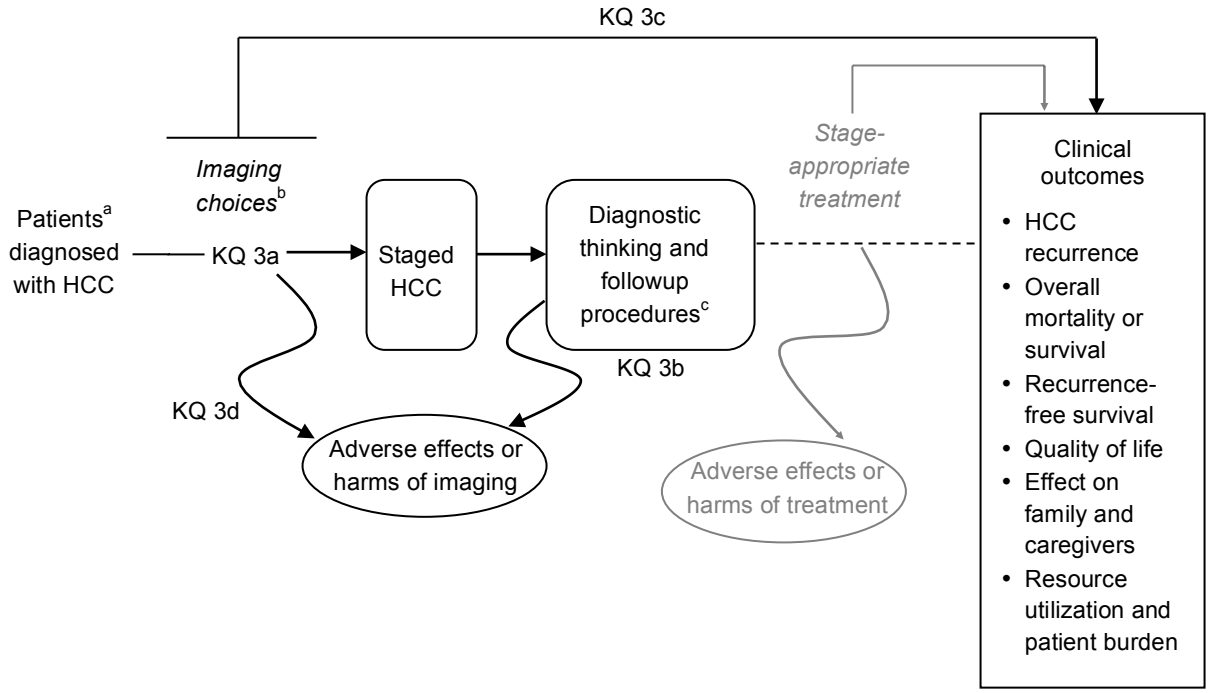


Figure 2. Analytic framework—diagnosis (Key Question 2)



[a] Potential modifiers of test performance include patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, and other factors (e.g., biomarker levels, setting).
[b] Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers
HCC = hepatocellular carcinoma; KQ = Key Question

Figure 3. Analytic framework—staging (Key Question 3)



[a] Potential modifiers of test performance include patient (e.g., severity of liver disease, underlying cause of liver disease, body mass index, age, sex, race), tumor (e.g., tumor diameter, degree of differentiation, location), technical, and other factors (e.g., biomarker levels, setting).
[b] Imaging techniques are used singly, in combination, or in sequence with or without biomarkers used as modifiers
[c] Followup procedures include biopsy.
HCC = hepatocellular carcinoma; KQ = Key Question

Table 1. Imaging techniques used in the surveillance, diagnosis, and staging of hepatocellular carcinoma

Imaging Modality	Key Characteristics	Surveillance	Diagnosis	Staging
Transabdominal ultrasound (US)	This modality uses ultrasound waves and their reflection from tissue interfaces to generate images of the underlying anatomy. Conventional (noncontrast) ultrasound is limited in its ability characterize hepatic lesions. Use of intravenous (IV) microbubble contrast agents has been proposed as a method for improving the characterization of liver masses. Most studies of contrast-enhanced US have focused on targeted evaluation of lesions identified on nonenhanced US or other imaging studies, due to the limited duration that contrast is present in the liver.	●	● (IV contrast only)	
Computed tomography (CT)	This imaging modality is based on x-ray exposure and acquisition of data through a set of detectors arrayed in a linear fashion. Contrast-enhanced CT images are obtained after injecting iodinated IV contrast media. Multiple passes are performed at specific times following contrast administration (multiphase contrast study). **Spiral CT** performs continuous scans, acquiring information to generate images in multiple planes. **Multidetector CT** scanners are based on the same imaging principles as spiral CT but utilize a two-dimensional array of detectors. MDCT permits faster scanning, resulting in fewer motion artifacts and improved image quality. **Dual energy CT** is a newer technique that uses x-rays of varying energy (70–140 kVp) to increase tissue contrast and detect different elements (e.g., iodine, calcium) within the liver. **Spectral CT** is a related technique that can separate and utilize information from the whole x-ray spectrum.	●	●	●
Magnetic resonance imaging (MRI)	This imaging technique uses a strong magnetic field and radiofrequency pulses to obtain anatomic images of the body. MRI scanning is slower than CT scanning and requires that the patient remain still during image acquisition. Like CT, multiphase MRI images are obtained in multiple passes following the IV administration of gadolinium-based contrast agents. MRI imaging acquisition techniques can preferentially assess tissues for fat content, diffusion characteristics, and edema. Different gadolinium contrast media are available, including nonspecific arterially enhancing agents such as gadopentetate and gadodiamide, and newer hepatic-specific agents like gadoxetic acid disodium or gadobenate that are preferentially taken up by functioning hepatocytes and excreted in the biliary system.	●	●	●
Positron emission tomography (PET)	This functional imaging technique uses radioisotope-tagged tracers to examine the level and type of biochemical activity in lesions suspected to be cancerous throughout the body (making it useful to study metastases). The most commonly used tracer is ^{18}F-fluorodeoxyglucose (FDG), which detects cells exhibiting increased glucose transport and metabolism (cancer cells typically exhibit such metabolic activity). Alternative tracers have also been investigated.			●

CT = computed tomography; FDG = ^{18}F-fluorodeoxyglucose; IV = intravenous; MRI = magnetic resonance imaging; PET = positron emission tomography; US = ultrasound

Table 2. Recommended minimum technical specifications for dynamic contrast-enhanced computed tomography (CT) of the liver[a]

Feature	Specification	Comment
Scanner type	Multi–detector row scanner	. . .
Detector type	Minimum of 8 detector rows	Need to be able to image entire liver during brief late arterial phase time window
Reconstructed section thickness	Minimum of 5-mm reconstructed section thickness	Thinner sections are preferable, especially if multiplanar reconstructions are obtained
Injector	Power injector, preferably dual-chamber injector with saline flush	Bolus tracking recommended
Contrast agent injection rate	Minimum, 3 mL/sec; better, 4–6 mL/sec with minimum of 300 mg iodine per milliliter or higher, for dose of 1.5 mL/kg of body weight	. . .
Mandatory dynamic phases during contrast-enhanced CT*	Late arterial phase, portal venous phase, and delayed phase	Artery fully enhanced, beginning contrast enhancement of portal vein; portal vein enhanced, peak liver parenchymal enhancement, beginning contrast enhancement of hepatic veins; variable appearance, >120 sec after initial injection of contrast agent
Dynamic phases (timing)	Bolus tracking or timing bolus recommended for accurate timing	. . .

*Comments describe typical hallmark image features

[a]Reprinted with permission of the copyright holder, Radiological Society of North America (RSNA), from Wald C, Russo MW, Heimbach JK, et al. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology. 2013 Feb;266(2):376-82.[15]

CT = computed tomography

Table 3. Recommended minimum technical specifications for dynamic contrast-enhanced MR imaging of the liver[a]

Feature	Specification	Comment
MR unit type	1.5-T or greater main magnetic field strength	Low-field-strength magnets not suitable
Coil type	Phased-array multichannel torso coil	Unless patient-related factors preclude use (eg, body habitus)
Minimum sequences	Nonenhanced and dynamic gadolinium-enhanced T1-weighted GRE sequence (3D preferable), T2-weighted (with and without fat saturation), T1-weighted in- and opposed-phase imaging	. . .
Injector	Dual-chamber power injector	Bolus tracking recommended
Contrast agent injection rate	For extracellular gadolinium chelate that does not have dominant biliary excretion, 2–3 mL/sec	Preferably resulting in vendor-recommended total dose
Mandatory dynamic phases at contrast-enhanced MR imaging*	Nonenhanced T1 weighted, late arterial phase, portal venous phase, delayed phase	For nonenhanced T1 weighted, do not change imaging parameters for contrast-enhanced imaging; for late arterial phase, artery fully enhanced, beginning contrast enhancement of portal vein; for portal venous phase, portal vein enhanced, peak liver parenchymal enhancement, beginning contrast enhancement of hepatic veins; for delayed phase, variable appearance, >120 sec after initial injection of contrast agent
Dynamic phases, timing	Use of a bolus-tracking method for timing contrast agent arrival for late arterial phase imaging is preferable; portal venous phase (35–55 sec after initiation of late arterial phase imaging); delayed phase (120–180 sec after initial contrast agent injection)	. . .
Section thickness	For dynamic series, 5 mm or less; for other imaging, 8 mm or less	. . .
Breath holding	Maximum length of series requiring breath hold should be about 20 sec, with a minimum matrix of 128 × 256	Compliance with breath-hold instructions is very important; technologists need to understand the importance of patient instruction before and during imaging

Note.—GRE = gradient echo, 3D = three-dimensional.

*Comments describe typical hallmark image features.

[a]Reprinted with permission of the copyright holder, Radiological Society of North America (RSNA), from Wald C, Russo MW, Heimbach JK, et al. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology. 2013 Feb;266(2):376-82.[15]

MR = magnetic resonance

Methods

The methods for this systematic review follow the methods suggested in the AHRQ Effective Health Care program methods guides.[24,25]

Topic Refinement and Review Protocol

This topic was selected for review based on a nomination from the Tufts Evidence-Based Practice Center (EPC) topic identification project, which included a set of draft proposed Key Questions. The Key Questions and scope were further developed with input from a Technical Expert Panel (TEP). The TEP provided high-level content and methodological guidance to the review process through involvement of clinicians and researchers with expertise in the diagnosis and management of liver diseases and cancers, radiologists, hepatologists, clinical outcomes researchers, and patient and payer representatives. TEP members disclosed all financial and other conflicts of interest prior to participation. The AHRQ Task Order Officer and the investigators reviewed the disclosures and determined that the panel members had no conflicts of interest that precluded participation.

Based on feedback from the TEP, the Key Questions and protocol were revised. The revised Key Questions were posted to the Agency for Healthcare Research and Quality (AHRQ) public Web site for a four-week period of public comment. No public comments were received, and the Key Questions were subsequently finalized. The protocol for this comparative effectiveness review (CER) was drafted and reviewed by the TEP; it is available from the AHRQ Web site where it was posted on July 25, 2013.[26] The protocol was also registered in the PROSPERO international database of prospectively registered systematic reviews in health and social care.[27]

Searching for the Evidence

For the primary literature, we searched Ovid MEDLINE®, Scopus, Evidence-Based Medicine Reviews (Ovid), the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, and the Health Technology Assessment Database from 1998 through March 2013 (see Table 4 for search strategy). We searched for unpublished studies in clinical trial registries (ClinicalTrials.gov, Current Controlled Trials, ClinicalStudyResults.org, and the WHO International Clinical Trials Registry Platform), regulatory documents (FDA Medical Devices Registration and Listing), and individual product Web sites. Scientific information packets (SIPs) were solicited via a notice published in the Federal Register that invited interested parties to submit relevant published and unpublished studies using the publicly accessible AHRQ Effective Health Care online SIP portal.[28] One SIP response was received, but it yielded no additional relevant studies. We also hand-searched the reference lists of relevant studies and previous systematic reviews for additional studies.

Populations and Conditions of Interest

The populations and conditions of interest for each key question are described in Table 5.

Study Selection

We developed criteria for inclusion and exclusion of studies based on the Key Questions and the populations, interventions, comparators, outcomes, timing, and setting (PICOTS) approach (Table 5). Papers were selected for full review if they were about imaging for HCC, were relevant to one or more Key Questions, and met the predefined inclusion criteria. We excluded studies published only as conference abstracts, restricted inclusion to English language articles, and excluded studies of nonhuman subjects. Studies had to report original data to be included.

Each abstract was reviewed for inclusion. Full-text articles were obtained for all studies that investigators identified as potentially meeting inclusion criteria. Two investigators independently reviewed all full-text articles for final inclusion or exclusion. Discrepancies were resolved through discussion and consensus.

We selected studies of adults undergoing surveillance for HCC (Key Question 1), imaging for further evaluation of a hepatic lesion or to distinguish HCC from another type of hepatic lesion (Key Question 2), and staging of HCC (Key Question 3). For Key Question 1, we also included studies that reported diagnostic accuracy of imaging for identification of HCC in nonsurveillance settings, such as series of patients undergoing liver transplantation, hepatic resection, or ablative therapy, or series of patients in whom the prevalence of HCC was known. Although imaging was not specifically performed for surveillance in these studies, they were reviewed under Key Question 1 because they were designed to assess the utility of imaging modalities for identifying HCC lesions, rather than to further characterize or assess a previously identified lesion (Key Question 2). We excluded studies that reported diagnostic accuracy of imaging for non-HCC malignant lesions, including metastatic lesions to the liver. We included studies that reported diagnostic accuracy for HCC and cholangiocarcinoma together only if the proportion of patients with cholangiocarcinoma was <10 percent.

We selected studies of US (with or without contrast enhancement), contrast-enhanced CT (nonmultidetector or multidetector spiral CT, and dual energy or spectral CT), contrast-enhanced MRI, and PET or PET/CT using various tracers. We excluded studies of nonspiral CT and MRI using machines ≤1.0 T, as these are considered outdated techniques.[15] We excluded studies published prior to 1998 and also excluded studies in which imaging commenced prior to 1995, unless those studies reported use of imaging meeting minimum technical criteria (defined as nonmultidetector or multidetector spiral CT and MRI with a 1.5 or 3.0 T machine). We excluded studies that evaluated MRI with agents that are no longer produced commercially and are unavailable for clinical use—superparamagnetic iron oxide (ferumoxides or ferucarbotran) or mangafodipir contrast—unless results based on gadolinium-enhanced imaging phases were reported separately. Although US microbubble contrast agents are not approved by the U.S. Food and Drug Administration (FDA) for evaluation of liver lesions, we included such studies because these agents are available commercially outside the United States, US with contrast is commonly performed in other countries (including in Europe), and efforts to obtain FDA approval are ongoing.[29-31] We excluded studies of CT arterial portography and CT hepatic angiography, which are invasive techniques not typically utilized in the United States for diagnosis and staging of HCC. We also excluded studies of intraoperative US.

For studies of test performance (e.g., sensitivity, specificity, and likelihood ratios), we included studies that evaluated one or more imaging methods against a reference standard. Reference standards were histopathology (based on explanted liver or nonexplant histological specimen from surgery or percutaneous biopsy), imaging and clinical followup, or some combination of these standards. We excluded studies in which the reference standard involved

the imaging test under evaluation and studies that had no reference standard (i.e., reported the number of lesions identified with an imaging technique but did not evaluate accuracy against a reference technique).

To assess comparative effects of imaging on clinical outcomes (e.g., mortality, HCC recurrence, quality of life, and harms), we included randomized controlled trials that compared different imaging modalities or strategies. A systematic review funded by the Department of Veterans Affairs Evidence Synthesis program on effects of screening for HCC on clinical outcomes is currently in progress that will also include comparative observational studies.[32]

To assess comparative effects of imaging on intermediate outcomes (e.g., effects on diagnostic thinking, clinical thinking, and resource utilization), we included randomized trials and cohort studies that compared different imaging modalities or strategies.

Data Abstraction and Data Management

We extracted the following data from included studies into evidence tables using Excel spreadsheets: study design, year, setting, country, sample size, method of data collection (retrospective or prospective), eligibility criteria, population and clinical characteristics (including age, sex, race, underlying cause of liver disease, proportion of patients in sample with HCC, HCC lesion size, and proportion with cirrhosis), the number of readers, criteria used for a positive test, and the reference standard used. We abstracted results for diagnostic accuracy, intermediate outcomes, and clinical outcomes, including results stratified according to patient, lesion, and imaging characteristics. Technical information for different imaging tests was abstracted as follows:[15]

- Ultrasound
  - Use of contrast
  - Type of contrast
  - Ultrasound operator (technician, physician, or other)
  - Transducer frequency
  - Use of Doppler
- Computed tomography
  - Use of multidetector scanner and the number of rows
  - Imaging sequences with timing
  - Contrast rate
  - Section thickness for contrast-enhanced images
  - Use of dual energy or spectral CT techniques
- Magnetic resonance imaging
  - MRI unit type (number of Teslas)
  - Imaging sequences with timing
  - Type of contrast
  - Section thickness
  - Use of diffusion-weighted imaging sequences
  - Spatial resolution
- Positron emission tomography
  - PET scanner versus PET/CT
  - Tracer type

Data abstraction for each study was completed by two investigators: the first abstracted the data, and the second reviewed the abstracted data for accuracy and completeness against the original articles. A team member with expertise in abdominal imaging reviewed data abstractions related to technical specifications.

Assessment of Methodological Risk of Bias of Individual Studies

We assessed risk of bias (quality) for each study based on predefined criteria. Randomized trials and cohort studies were evaluated using criteria and methods developed by the U.S. Preventive Services Task Force.[33] These criteria were applied in conjunction with the approach recommended in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.[24] Studies of diagnostic test performance were assessed using the approach recommended in the AHRQ Methods Guide for Medical Test Reviews,[25] which is based on methods developed by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) group.[34]

Individual studies were rated as having "low," "moderate," or "high" risk of bias. Studies rated "low" risk of bias are generally considered valid. Randomized trials and cohort studies assessed as having low risk of bias have a clear description of the population, setting, interventions, and comparison groups; a valid method for allocating patients to treatment (for randomized trials); clear reporting of dropouts with low dropout rates; appropriate methods for preventing bias; appropriate measurement of and analysis of confounders (for cohort studies); and appropriate measurement of outcomes. Studies of diagnostic test performance that are assessed as having low risk of bias use unbiased methods to enroll patients, avoid use of a case-control design, use a credible reference standard, apply the same reference standard to all patients, use blinded interpretation of the diagnostic test as well as the reference standard, use preset criteria to define a positive test, avoid long delays between the imaging test and the reference standard, and have limited (defined for this report as <10%) loss to followup.[7,12] We considered studies in which all patients had HCC as utilizing a case-control design, even if some patients also had other (non-HCC) lesions. We considered studies that utilized a histopathological reference standard or a reference standard consistent with EASL or AASLD criteria to be adequate. The AASLD criteria is based on tumor size, with lesions <1 cm undergoing serial imaging followup.[9] Based on AASLD criteria, for lesions >1 cm, presence of a typical enhancement pattern on CT or MRI is considered diagnostic for HCC; for lesions without typical enhancement on one of these imaging tests, biopsy is required.

Studies rated as having "moderate" risk have some methodological shortcomings, but no flaw or combination of flaws judged likely to cause major bias. In some cases, the study may be missing information, making it difficult to assess its methods or potential limitations. The moderate risk of bias category is broad, and studies with this rating will vary in their strengths and weaknesses; the results of some studies assessed to have moderate risk of bias are likely to be valid, while others may be only possibly valid.

Studies rated as having "high" risk of bias have significant flaws that may invalidate the results. They have a serious or "fatal" flaw in design, analysis, or reporting; large amounts of missing information; or serious discrepancies in reporting. The results of these studies are at least as likely to reflect flaws in the study design as the differences between the compared interventions. We did not exclude studies rated as having high risk of bias a priori, but they were

considered the least reliable when synthesizing the evidence, particularly when discrepancies between studies were present.

Data Synthesis

We performed meta-analyses on measures of test performance in order to help summarize data and obtain more precise estimates.[35] We only pooled studies that were clinically comparable and could provide a meaningful combined estimate (based on the variability among studies in design, patient population, imaging methods, and outcomes) and magnitude of effect size. We conducted separate analyses for each imaging modality, stratified according to the unit of analysis used (patients with HCC, HCC lesions, or liver segments with HCC). For studies that used multiple readers, we averaged results across readers. For Key Question 1, we also stratified analyses according to whether imaging was performed for surveillance or if imaging was performed in a series of patients for some other reason. For Key Question 2, we separately analyzed studies that evaluated further imaging of a focal liver lesion identified on previous imaging and studies evaluating the ability of imaging tests to distinguish between HCC and another specific non-HCC lesion. For Key Question 3, we separately analyzed studies on test performance of imaging for identifying metastatic HCC and accuracy of imaging for staging based on tumor, node, metastasis staging (TNM), Barcelona Liver Cancer Clinic (BLCC), and other criteria.

We evaluated a number of potential sources of heterogeneity (see below) and modifiers of diagnostic accuracy. We performed analyses stratified according to the reference standard used and on domains related to risk of bias, aspects of study design (retrospective or prospective, use of a confidence rating scale), setting (based on country in which imaging was performed), and technical factors (such as scanner types, type of contrast or tracer used, use of recommended imaging phases, timing of delayed phase imaging, and section thickness). We also evaluated diagnostic accuracy in subgroups stratified according to HCC lesion size, degree of tumor differentiation, and tumor location, as well as patient characteristics such as severity of underlying liver disease, underlying cause of liver disease, and body mass index. For analyzing effects of tumor size and degree of differentiation on estimates of accuracy, we analyzed studies on surveillance and diagnosis together.

We performed separate analyses on the subset of studies that directly compared two or more imaging modalities or techniques in the same population against a common reference standard. Research indicates that results based on such direct comparisons differ from results based on noncomparative studies, and may be better suited for evaluating comparative diagnostic test performance.[36]

We did not perform meta-analysis on staging accuracy and intermediate or clinical outcomes due to the small number of studies. Rather, we synthesized these studies qualitatively, using the methods described below for assessing the strength of evidence.

Approaches to Data Analysis

We conducted meta-analysis to quantitatively synthesize data and obtain summary estimates of test performance. We used a bivariate logistic mixed effects model[37] to analyze sensitivity and specificity, incorporating the correlation between sensitivity and specificity. We assumed random effects across studies for sensitivity and specificity, and heterogeneity among the studies was measured based on the random effect variance (τ^2). The advantage of using a logistic mixed

effects model is that it handles sparse data better and does not need to assume an ad hoc continuity correction when a study has zero events.[37] We calculated positive likelihood ratio (LR+) and negative likelihood ratio (LR-) using the summarized sensitivity and specificity.[38,39]

The data were synthesized by each imaging modality. To address possible source of heterogeneity among studies and produce meaningful summary estimates, we stratified analyses by Key Questions and the unit of analysis. We also conducted extensive subgroup and sensitivity analyses based on study-level, patient, tumor, technical, and other factors. When data were available, we performed separate meta-analyses for within-study comparisons based on technical factors, lesion size, and degree of tumor differentiation.

To assess the comparative effectiveness of alternative imaging modalities, we also conducted head-to-head comparisons between imaging modalities when data were available, using the same bivariate logistic mixed effects model as described above, but adding an indicator variable for imaging modalities (equivalent to a meta-regression approach). These analyses only included studies that directly compared two imaging modalities, in order to restrict the comparison to direct evidence. Again we stratified the comparisons by Key Questions and unit of analysis, and we conducted subgroup analyses by methodological and lesion characteristics when the data allowed. All analyses were conducted using SAS 9.3 (SAS institute Inc., Cary, NC).

Grading the Strength of Evidence for Individual Comparisons and Outcomes

The strength of evidence for each Key Question was assessed by one researcher for each outcome described in the PICOTS using the approach described in the AHRQ Methods Guide.[24] The strength of evidence was based on the overall quality of each body of evidence, based on the risk of bias (graded low, moderate, or high); the consistency of results between studies (graded consistent, inconsistent, or unknown/not applicable when only one study was available); the directness of the evidence linking the intervention and health outcomes (graded direct or indirect); and the precision of the estimate of effect, based on the number and size of studies and confidence intervals for the estimates (graded precise or imprecise. We did not downgrade a body of evidence for directness that evaluated an intermediate outcome, if the intermediate outcome (such as diagnostic accuracy or effects on diagnostic thinking) was the specific focus of the Key Question. We did not grade supplemental domains for cohort studies evaluating intermediate and clinical outcomes because too few studies were available for these factors to impact the strength of evidence grades. We did not assess studies of diagnostic test performance for publication bias using graphical or statistical methods because research indicates that such methods can be very misleading. Rather, we searched for unpublished studies through searches of clinical trials registries and regulatory documents and by soliciting SIPs.

We graded the strength of evidence for each Key Question using the four key categories recommended in the AHRQ Methods Guide.[24] A “high” grade indicates high confidence that the evidence reflects the true effect and that further research is very unlikely to change our confidence in the estimate of effect. A “moderate” grade indicates moderate confidence that the evidence reflects the true effect and further research may change our confidence in the estimate of effect and may change the estimate. A “low” grade indicates low confidence that the evidence reflects the true effect and further research is likely to change the confidence in the estimate of effect and is likely to change the estimate. An “insufficient” grade indicates evidence either is unavailable or is too limited to permit any conclusion.

Assessing Applicability

We recorded factors important for understanding the applicability of studies, such as whether the publication adequately described the study population, the country in which the study was conducted, the prevalence of HCC in the patients who underwent imaging, the magnitude of differences in measures of diagnostic accuracy and clinical outcomes, and whether the imaging techniques were reasonably representative of standard practice.[40] We also recorded the funding source and role of the sponsor. We did not assign a rating of applicability (such as high or low) because applicability may differ based on the user of the report.

Peer Review and Public Commentary

Experts in gastroenterology, hepatology, and radiology, along with individuals representing stakeholder and user communities, will be invited to provide external peer review of a draft of this CER; AHRQ and an EPC associate editor will also provide comments. The draft report will be posted on the AHRQ Web site for 4 weeks to elicit public comment. All comments will be reviewed and addressed in a disposition of comments report that will be made available 3 months after the Agency posts the final CER on the AHRQ Web site (www.effectivehealthcare.ahrq.gov).

.

Table 4. Search strategy—Ovid MEDLINE® (1998–2013)

Search Strategy for Hepatocellular Carcinoma Imaging

1	Carcinoma, Hepatocellular/
2	Liver Neoplasms/
3	("hepatocellular cancer" or "hepatocellular carcinoma" or "HCC").ti,ab.
4	Diagnostic Imaging/
5	Ultrasonography/
6	Magnetic Resonance Imaging/
7	exp Tomography, Emission-Computed/ or exp Positron-Emission Tomography/ or exp Tomography, Spiral Computed/
8	("CT" or "dynamic multidetector computed tomography" or "MDCT" or "spiral CT" or "dual source CT" or "contrast CT" or "MRI" or "FDG-PET").ti,ab.
9	or/1-3
10	or/4-8
11	9 and 10
12	"Sensitivity and Specificity"/
13	"Predictive Value of Tests"/
14	ROC Curve/
15	"Reproducibility of Results"/
16	(sensitiv$ or "predictive value" or accurac$).ti,ab.
17	or/12-16
18	11 and 17
19	limit 18 to yr="1998 - 2013"

Table 5. Inclusion and exclusion criteria by Key Question

	Inclusion Criteria	Exclusion Criteria
All Key Questions		
Interventions	• Ultrasound (including noncontrast enhanced and contrast enhanced) • Contrast-enhanced spiral CT (including nonmultidetector or multidetector CT, and CT using dual energy and spectral methods) • Contrast-enhanced MRI using a 1.5 or 3.0 T scanner • Diffusion-weighted MRI • PET or PET/CT (including use of ^{18}F-fluorodeoxyglucose , ^{18}F-fluorothymidine, ^{11}C-acetate, and ^{11}C-choline tracers)	• Outdated imaging techniques (e.g., conventional, nonspiral/nonmultidetector CT, MRI using a ≤1.0 T scanner) • CT and MRI without contrast, with the exception of studies of diffusion-weighted MRI without contrast • CT arterial portography and CT hepatic arteriography • C-arm CT • Intraoperative ultrasound • MRI with ferucarbotran, ferumoxides, or mangafodipir contrast
Comparisons	For studies of diagnostic accuracy (comparative test performance): • Reference standard comparators: Histopathology (based on explanted liver specimens or biopsy), clinical followup, and imaging followup • Imaging comparators: Alternative imaging tests or strategies. For studies of comparative effectiveness: • No imaging or an alternative imaging strategy	Does not meet inclusion criteria
Timing	No restrictions	None
Setting	All care settings (e.g., primary and secondary care)	None
Study Designs	• Controlled randomized and nonrandomized trials • Cohort studies on effects of imaging on diagnostic thinking or clinical decisionmaking • Studies of diagnostic accuracy	• Studies of diagnostic accuracy that did not report the reference standard used, or in which the reference standard included the results of the test being investigated • Case reports, case series, letters to the editor, and nonsystematic reviews • Studies published prior to 1998 or in which imaging was performed prior to 1995, unless technical details were reported and studies met minimum technical criteria as described in the Interventions section above
Key Questions 1 and 2		
Populations	Key Question 1 • Patients at high risk for HCC undergoing surveillance, including: Asian male HBV carriers over age 40, Asian female HBV carriers over age 50, HBV carriers with a family history of HCC, African/North American black HBV carriers, all individuals with cirrhosis (including alcoholic cirrhosis), HBV or HCV carriers with cirrhosis, and patients with stage 4 primary biliary cirrhosis • Other high-risk patients undergoing surveillance as defined by the primary studies • Patients enrolled in studies designed to determine detection rates of imaging for HCC, including patients who underwent liver transplantation or surgery for HCC or other reasons. Key Question 2 • Patients in whom a suspicious lesion(s) for HCC has been detected by surveillance or by other means, including patients who underwent liver transplantation for HCC or other reasons	• Patients with cholangiocarcinoma, unless they comprised <10% of the study population • Patients with nonprimary (metastatic) lesions to the liver • Patients undergoing imaging to evaluate response to ablative or other treatments • Children.

	Inclusion Criteria	Exclusion Criteria
	• Patients enrolled in studies designed to distinguish HCC from another type of liver lesion (benign or malignant).	
Outcomes	• Diagnostic outcomes: test performance, types of HCC lesions detected • Intermediate outcomes: effects on diagnostic thinking, effects on clinical decisionmaking. • Clinical and patient-centered outcomes: overall mortality or survival, recurrence of HCC, including rates of seeding by fine-needle aspiration; quality of life as measured with scales such as the Short-Form Health Survey or EuroQol 5D; and psychosocial effects of diagnostic testing on patients, patients' caregivers, and other family members • Resource utilization and patient burden (e.g., costs associated with the imaging procedure, the number of imaging procedures, and other procedures conducted) • Harms: adverse effects or harms associated with the imaging techniques (e.g., test-related anxiety, adverse events secondary to venipuncture, contrast allergy, exposure to radiation) and adverse effects or harms associated with test-associated diagnostic workup (e.g., harms of biopsy or harms associated with workup of other incidental tumors discovered on imaging).	• Nonclinical and nondiagnostic outcomes.
Key Question 3		
Populations	• Patients diagnosed with HCC undergoing staging before initial treatment.	• Patients with cholangiocarcinoma, unless they comprised <10% of the study population • Patients with nonprimary (metastatic) lesions to the liver • Patients undergoing imaging to evaluate response to ablative or other treatments • Children.
Outcomes	• Measures of stage-specific accuracy of imaging (e.g., proportion correctly staged, understaged, and overstaged) • Intermediate and clinical outcomes as described for Key Questions 1 and 2.	• Nonclinical and nondiagnostic outcomes.

CT = computed tomography; HBV = hepatitis B virus; HCC = hepatocellular carcinoma; HCV = hepatitis C virus; MRI = magnetic resonance imaging; PET = positron emission tomography

Results

Introduction

The bulk of the available evidence addresses diagnostic accuracy of imaging techniques for hepatocellular carcinoma (HCC)—ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). Very few studies compared effects of different imaging modalities or strategies on diagnostic thinking, clinical outcomes, and almost no studies reported harms.

Results of Literature Searches

The search and selection of articles are summarized in the study flow diagram (Figure 4). Of the 4476 citations identified at the title and abstract level, 759 articles appeared to meet inclusion criteria and were selected for further full-text review. Of the 759 articles reviewed at the full-text level, a total of 255 studies met inclusion criteria (Appendix A); primary reasons for exclusion of the articles reviewed at the full-text level were (Appendix B). We rated three studies low risk of bias,[41-43] 164 moderate risk of bias, and 88 high risk of bias (Appendix C).

One cohort study[44] and three randomized trials reported patient outcomes of imaging for staging.[45-47] We identified 251 studies that evaluated diagnostic accuracy of imaging tests.

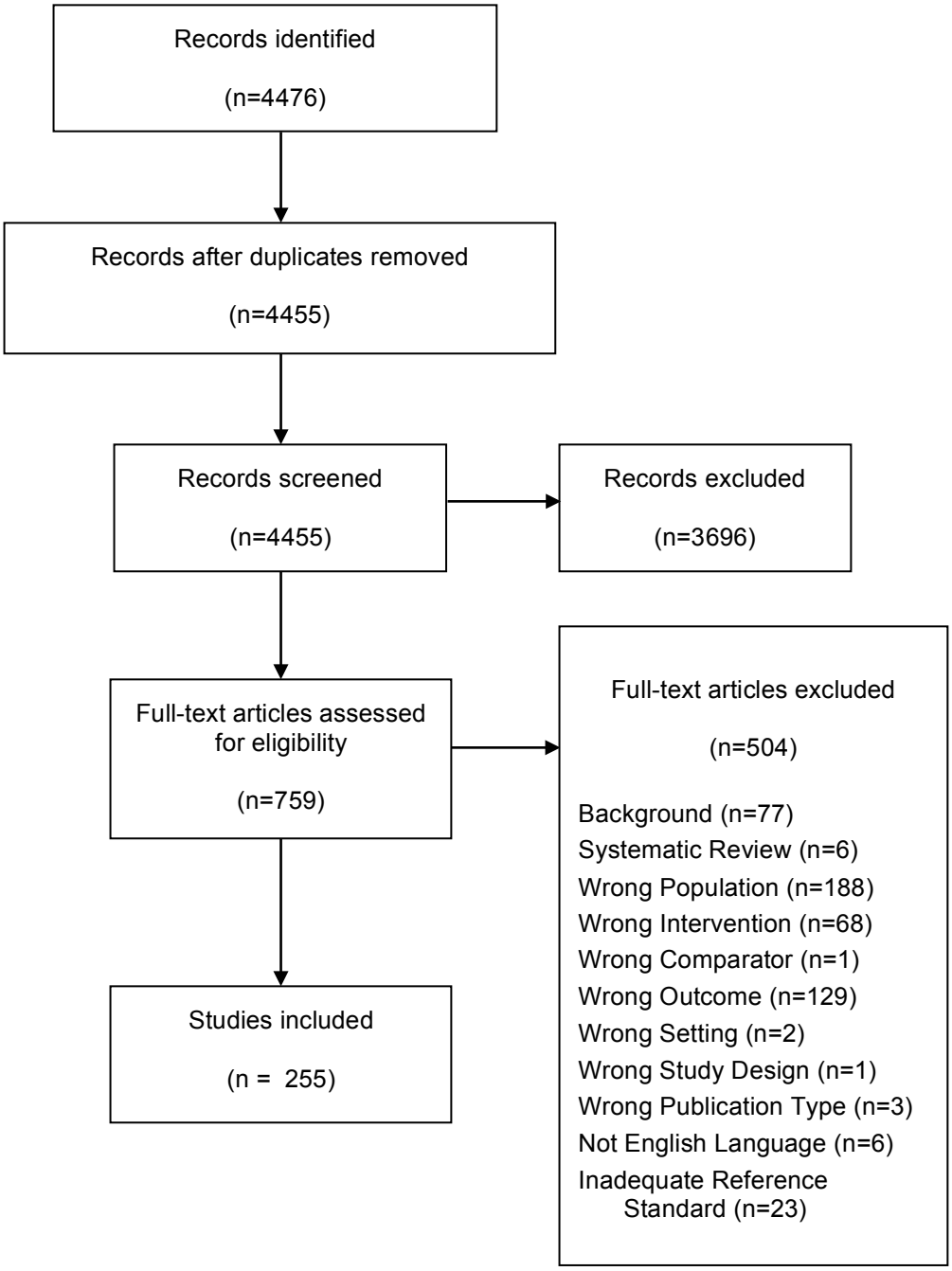
Of the diagnostic accuracy studies, 60 evaluated US (Appendix D),[42,48-106] 125 evaluated CT (Appendix E),[41-43,48,50-53,56,57,61,64,68,69,75,79,81,83,87,89,91-93,98-100,104-202] 117 evaluated MRI (Appendix F),[42,43,48,53,55,56,60,64,68,69,75,89,90,93,95,98,105,108,110,113,116,118,121-123,125,127,128,131,132,143,148,149,151,153,161,164,169,175,176,178,180,183,185,188,191,193,195,200,203-270] and 31 evaluated PET (Appendix G).[98,99,114,137,139,156,174,271-294] Twenty-eight studies evaluated use of more than one imaging technique in combination or sequentially.[42,43,48,52,55,56,64,69,73,87,114,121,137,156,174,271,272,275,277,279,283,284,286-288,291,294,295] Almost all studies reported sensitivity, but only 117 reported specificity or provided data to calculate specificity. We found 141 studies avoided use of a case-control design, 137 used blinded ascertainment, and 69 used a prospective design. More studies were conducted in Asia (163 studies) than in the United States or Europe (86 studies). In 136 studies, imaging was conducted starting in or after 2003.

Twenty-four studies evaluated CT using methods that met minimum technical specifications (≥8-row multidetector CT; contrast rate ≥3 ml/s; at least arterial, portal venous, and delayed phase imaging; delayed phase imaging performed >120 s following administration of contrast; and enhanced imaging section thickness ≤5 mm)[41,42,53,64,69,79,87,118,122,128,143,147-149,153,162,163,168,169,180,185,191,194,197] and 50 studies evaluated MRI using methods that met minimum technical specifications (1.5 or 3.0 T MRI; at least arterial, portal venous, and delayed phase imaging; delayed phase imaging performed >120 s following administration of contrast; and enhanced imaging section thickness ≤5 mm).[42,53,55,64,68,95,108,110,118,121,122,128,132,143,147,148,153,169,176,180,183,188,191,203,205,206,208,217,220,222-234,240-243,249,251,255,259] 52 MRI studies evaluated use of hepatic-specific contrast (e.g., gadoxetic acid or gadobenate).[42,48,53,60,64,68,69,73,95,108,110,118,122,123,128,131,132,143,148,153,169,176,180,188,191,203-206,208,211,220,221,223,225-230,232,233,242,243,247-249,251,255,259,264,295] 36 US studies evaluated use of microbubble contrast agents.[42,48,50,52,54,55,58,60-64,66,68,69,72,73,78-81,83,84,86-88,90,91,94-97,101-104] 29 studies evaluated PET using FDG,[98,99,114,137,139,156,174,271-273,275-279,281-294] eight studies using ^{11}C-

acetate,[114,272,275-277,280,283,293] and three studies evaluated use of other tracers (^{18}F-fluorothymidine or ^{18}F-fluorocholine).[274,287,288]

Data for outcomes other than measures of test performance were sparse. Seven studies reported comparative effects on diagnostic thinking,[75,92,109,114,164,187,192] three studies reported comparative clinical and patient-centered outcomes,[45-47] and three studies reported harms associated with imaging for HCC.[83,131,189]

Figure 4. Study flow diagram



Key Question 1. What is the comparative effectiveness of available imaging-based surveillance strategies for detecting HCC among individuals undergoing surveillance for HCC?

Description of Included Studies

Six studies[51,57,82,85,93,100] evaluated diagnostic accuracy of imaging techniques for surveillance and 174 studies[41,48,49,53,56,59,62,64-66,68,70-73,75-77,89,98,99,105-108,110-114,116-120,122-128,130-136,138-140,142-155,157-165,167-173,175-190,192-195,197-208,213-217,220,223,225,227-242,245,246,248-254,257,258,260-263,265,267,268,270,272,274,276,277,280-284,287-293,295] reported diagnostic accuracy in nonsurveillance settings (e.g., imaging performed to assess detection rates in a series of patients undergoing treatment for HCC or patients with otherwise known prevalence of HCC prior to imaging). Four studies of PET evaluated accuracy specifically for identification of recurrent HCC.[271,279,286,292]

One randomized trial (rated high risk of bias)[47] evaluated clinical outcomes associated with imaging-based surveillance versus no screening, and two trials[45,46] evaluated clinical outcomes associated with different US surveillance intervals. No study compared effects of different imaging surveillance strategies on diagnostic thinking or clinical decisionmaking. Two studies reported harms associated with imaging for HCC.[131,189]

Key Points

Test performance

- For surveillance, using patients with HCC as the unit of analysis:
  - US without contrast: Sensitivity was 0.82 (95% CI 0.66 to 0.092, 3 studies) and specificity was 0.87 (95% CI 0.77 to 0.93, 2 studies), for a LR+ of 6.2 (95% CI 3.6 to 11) and LR- of 0.20 (0.10-0.40).
  - CT: Sensitivity was 0.84 (95% CI 0.59 to 0.95, 2 studies) and specificity 0.99 (95% CI 0.86 to 0.999, 2 studies).
  - MRI and PET were not evaluated in surveillance settings.
- For surveillance, using HCC lesions as the unit of analysis:
  - US without contrast: Sensitivity was 0.60 (95% CI 0.36 to 0l.80, 2 studies) and specificity was 0.94 (95% CI 0.83 to 0.98, 1 study), for a LR+ of 9.8 (95% CI 3.7 to 26) and LR- of 0.43 (95% CI 0.24 to 0.74.
  - CT: Sensitivity was 0.62 (95% CI 0.46 to 0.76, 1 study).
- In nonsurveillance settings (e.g., imaging in series of patients who underwent liver transplantation or resection, or series of patients in whom the prevalence of HCC is known), using patients with HCC as the unit of analysis:
  - US without contrast: Sensitivity was 0.73 (95% CI 0.46 to 0.90, 8 studies) and specificity was 0.93 (95% CI 0.85 to 0.97, 6 studies), for a LR+ of 11 (95% CI 5.4 to 21) and LR- of 0.29 (95% CI 0.13 to 0.65).
  - CT: Sensitivity was 0.83 (95% CI 0.75 to 0.89, 16 studies) and specificity was 0.92 (95% CI 0.86 to 0.96, 11 studies), for a LR+ of 11 (95% CI 5.6 to 20) and LR- of 0.19 (95% CI 0.12 to 0.28).

- MRI: Sensitivity was 0.87 (95% CI 0.77 to 0.93, 11 studies) and specificity was 0.88 (95% CI 0.79 to 0.93, 9 studies), for a LR+ of 7.2 (95% CI 3.9 to 13) and LR- of 0.15 (95% CI 0.08 to 0.27).
- PET: For FDG PET, sensitivity was 0.52 (95% CI 0.39 to 0.66, 15 studies) and specificity was 0.95 (95% CI 0.92 to 0.99, 5 studies), for a LR+ of 11 (95% CI 2.6 to 49) and LR- of 0.50 (955 CI 0.37 to 0.68). For ^{11}C-acetate PET or PET/CT, sensitivity was 0.85 (95% CI 0.67 to 0.94, 4 studies). Specificity was not reported.

- In nonsurveillance settings, using HCC lesions as the unit of analysis:
  - US without contrast: Sensitivity was 0.60 (95% CI 0.42 to 0.75, 11 studies). Only two studies reported specificity, with inconsistent results (0.63, 95% CI 0.53 to 0.73 and 0.95, 95% CI 0.85 to 0.99).
  - US with contrast: Sensitivity was 0.76 (95% CI 0.53 to 0.90, 6 studies). No study evaluated specificity.
  - CT: Sensitivity was 0.77 (95% CI 0.73 to 0.81, 75 studies) and specificity was 0.89 (95% CI 0.83 to 0.93, 20 studies), for a LR+ of 7.0 (95% CI 4.6 to 11) and LR- of 0.25 (95% CI 0.21 to 0.30).
  - MRI: Sensitivity was 0.83 (95% CI 0.80-0.86, 68 studies) and specificity was 0.83 (95% CI 0.70-0.92, 13 studies), for a LR+ of 5.0 and LR- of 0.20 (95% CI 0.16 to 0.2.6).
  - PET: For FDG PET, sensitivity was 0.56 (95% CI 0.41 to 0.69, 4 studies) and specificity was 0.91 (95% CI 0.76 to 0.98, 1 study). For ^{11}C-acetate PET, sensitivity was 0.78 (95% CI 0.61 to 0.89, 4 studies). Specificity was not reported.

- Direct (within-study) comparisons of imaging modalities, using patients with HCC as the unit of analysis:
  - US without contrast versus CT: Sensitivity was 0.68 (95% CI 0.54 to 0.80) versus 0.80 (95% CI 0.68 to 0.88), for a difference of -0.12 (95% CI -0.20 to -0.03), based on 6 studies.
  - US without contrast versus MRI: Sensitivity was 0.61 (95% CI 0.48 to 0.74) versus 0.81 (95% CI 0.69 to 0.89), for a difference of -0.19 (95% CI -0.30 to -0.08), based on three studies.
  - MRI versus CT: Sensitivity was 0.88 (95% CI 0.53 to 0.98) versus 0.82 (95% CI 0.41 to 0.97), for a difference of 0.06 (95% CI -0.05 to 0.17), based on four studies.

- Direct (within-study) comparisons of imaging modalities, using HCC lesions as the unit of analysis:
  - US without contrast versus CT: Sensitivity was 0.55 (95% CI 0.43 to 0.66) versus 0.66 (95% CI 0.54 to 0.76) for a difference of -0.11 (95% CI -0.18 to -0.04), based on three studies.
  - US without contrast versus MRI: Sensitivity was 0.57 (95% CI 0.42 to 0.71) versus 0.79 (95% CI 0.67 to 0.88), for a difference of -0.22 (95% -0.31 to -0.14), based on three studies.
  - US with contrast versus CT: Sensitivity was 0.58 (95% CI 0.37 to 0.77 versus 0.74 (95% CI 0.54 to 0.87), for a difference of -0.16 (95% CI -0.32 to -0.01), based on three studies.

- US with contrast versus MRI: Sensitivity was 0.54 (95% CI 0.25 to 0.80) versus 0.70 (95% CI 0.40-0.89), for a difference of -0.16 (95% CI -0.30 to -0.02), based on 2 studies.
- MRI versus CT: Sensitivity was 0.81 (95% CI 0.77 to 0.84) versus 0.72 (95% CI 0.67 to 0.77), for a difference of 0.09 (95% CI 0.06 to 0.12), based on 28 studies. Findings were similar when studies were stratified according to use of nonhepatic-specific or hepatic-specific contrast.

- Multiple imaging modalities
  - One study found sensitivity of imaging with various combinations of two imaging modalities was similar or lower than single modality imaging, based on concordant positive findings on two imaging modalities. The other study reported higher sensitivity with multiple imaging modalities than with single modality imaging, but criteria for positive results based on multiple imaging modalities were unclear

- Sensitivity of US, CT, and MRI was lower in studies that used explanted liver as the reference standard than in studies that used other histopathological reference standards, clinical or imaging criteria, or a mixed reference standard.
  - US: Using HCC lesions as the unit of analysis, sensitivity was 0.34 (95% CI 0.21 to 0.49) in 5 studies that used explanted liver as the reference standard and ranged from 0.70 to 0.85 in studies that used other reference standards.
  - CT: Using HCC lesions as the unit of analysis, sensitivity was 0.69 (95% CI 0.60-0.77) in 21 studies that used explanted liver as the reference standard and ranged from 0.79 to 0.85 in studies that used other reference standards.
  - MRI: Using HCC lesions as the unit of analysis, sensitivity was 0.69 (95% CI 0.59 to 0.77) in 15 studies that used explanted liver as the reference standard and ranged from 0.85 to 0.88 in studies that used other reference standards.
  - PET: No study of FDG PET used an explanted liver reference standard. Three of the four studies that used HCC lesions as the unit of analysis used a nonexplant histological reference standard.
  - Specificity was reported in too few studies to draw strong conclusions.

- Across imaging modalities, based on within-study comparisons, sensitivity increased as HCC lesion size increased.
  - US: Sensitivity was 0.88 (95% CI 0.78 to 0.94) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions <2 cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies. The difference was larger in studies of US without contrast than studies of US with contrast, but these findings are difficult to interpret because sensitivity for HCC lesions <20 mm was much lower in the studies of US without contrast. For US without contrast, sensitivity was 0.09 (95% CI 0.02 to 0.29, 4 studies) for lesions <10 mm to 0.50 (95% CI 0.23 to 0.78, 4 studies) for lesions 10-20 mm and 0.88 (95% CI 0.66 to 0.96, 4 studies) for lesions >20 mm , for a difference of 0.37 (95% CI 0.18 to 0.57) for lesions >20 mm versus 10-20 mm, and 0.41 (95% CI 0.19 to 0.63) for lesions 10-20 mm versus <10 mm. For US with contrast, three studies found sensitivity of 0.64 (95% CI 0.33 to 0.87) for lesions 10-20 mm and 0.91 (95% CI 0.71 to 0.98) for lesions >20 mm, for a difference of 0.26 (95% CI 0.04 to 0.48).

- CT: Sensitivity was 0.94 (95% CI 0.91 to 0.95) for lesions >2 cm and 0.62 (95% CI 0.56 to 0.68) for lesions <2 cm, for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36), based on 33 studies. Sensitivity was 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10-20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 versus 10-20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10-20 versus <10 mm.
- MRI: Sensitivity was 0.96 (95% CI 0.93 to 0.97) for lesions >2 cm and 0.65 (95% CI 0.57 to 0.73) for lesions <2 cm, for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37), based on 25 studies. Sensitivity was 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10-20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 versus 10-20 mm and 0.34 (95% CI 0.27 to 0.41) for 10-20 versus <10 mm.
- PET: For FDG PET, sensitivity was consistently higher for larger lesions, based on five studies. Data were not pooled due to differences in the tumor size categories evaluated. Two studies of ^{11}C-acetatate PET found inconsistent effects of lesion size on sensitivity.

- Across imaging modalities, based on within-study comparisons, sensitivity was higher for moderately- or poorly-differentiated HCC lesions than for well-differentiated HCC lesions.
  - US with contrast: Sensitivity was 0.83 (95% CI 0.55 to 0.95) for moderately- or poorly-differentiated HCC lesions and 0.43 (95% CI 0.15 to 0.76) for well-differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.17 to 0.64), based on three studies.
  - CT: Sensitivity was 0.82 (95% CI 0.66 to 0.91) for moderately- or poorly-differentiated HCC lesions and 0.50 (95% CI 0.29 to 0.70) for well-differentiated lesions, for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45), based on five studies.
  - MRI: Sensitivity was 0.54 (95% CI 0.26 to 0.79) for moderately- or poorly-differentiated HCC lesions and 0.38 (95% CI 0.17 to 0.64) for well-differentiated lesions, but the difference was not statistically significant (0.16, 95% CI -0.11 to 0.43), based on two studies.
  - PET: For FDG PET, sensitivity was consistently higher for more poorly-differentiated lesions than more well-differentiated lesions, based on five studies. In three studies of ^{11}C-acetate PET and one study of ^{18}F-fluorochlorine, sensitivity for more well-differentiated lesions was not lower than for more poorly-differentiated lesions.

- Effects of other factors on estimates of test performance
  - US: In two studies that directly compared US with contrast versus without contrast, there was no clear difference in sensitivity (-0.04, 95% CI -0.11 to 0.04). One study that directly compared use of Doppler versus no Doppler showed no clear effect on estimates of sensitivity. Lesion depth and body mass index had no effect on estimates of sensitivity.

- CT: Using patients with HCC as the unit of analysis, studies with a contrast rate ≥3 ml/s reported a higher sensitivity (0.87, 95% CI 0.77 to 0.93, 8 studies) than studies with a contrast rate <3 ml/s (0.71, 95% CI 0.50-0.85, 4 studies). Studies with delayed phase imaging reported somewhat higher sensitivity (0.89, 95% CI 0.81 to 0.94, 7 studies) than studies without delayed phase imaging (0.74, 95% CI 0.66 to 0.87, 7 studies), but there were no clear effects in studies that used HCC lesions as the unit of analysis.
- MRI: There were no clear differences in estimates of diagnostic accuracy when studies were stratified according to MRI scanner type (1.5 vs. 3.0 T), imaging phases evaluated (with or without delayed phase imaging), timing of delayed phase imaging (>120 seconds vs. <120 seconds), section thickness (≤5 mm for enhanced images vs. >5 mm), or use of diffusion-weighted imaging. In studies that directly compared diagnostic accuracy with different types of contrast, hepatic-specific contrast agents (gadoxetic acid or gadobenate) were associated with slightly higher sensitivity than nonhepatic-specific contrast agents (gadopentetate or gadodiamide) (0.82, 95% CI 0.71 to 0.90 vs. 0.75, 95% CI 0.61 to 0.85, difference 0.07, 95% CI 0.01 to 0.14, 5 studies).
- PET: FDG PET was associated with lower sensitivity that ^{11}C-acetate PET when either patients (0.58 vs. 0.81, for a difference of -0.23, 95% CI -0.34 to -0.13, 3 studies) or HCC lesions (0.52 vs. 0.79, for a difference of -0.27, 95 %CI -0.36 to -0.17, 3 studies) were the unit of analysis. FDG PET was also associated with lower sensitivity than dual tracer PET with FDG and ^{11}C-acetate or ^{18}F-choline PET, but evidence was limited to 1 or 2 studies for each of these comparisons. Using patients as the unit of analysis, sensitivity of FDG PET (0.39, 95% CI 0.24 to 0.56, 8 studies) was lower than sensitivity of FDG PET/CT (0.65, 95% CI 0.50-0.78, 7 studies).

Diagnostic Thinking

- No study compared effects of different imaging modalities or strategies on diagnostic thinking, subsequent procedures, or resource utilization.

Clinical and Patient-centered Outcomes

- One cluster randomized controlled trial (n=18816) conducted in China found screening every 6 months with noncontrast US plus serum AFP versus no screening in persons 35 to 79 years of age (mean 42 years) with HBV infection or chronic hepatitis without HBV infection was associated with lower risk of HCC-related mortality (32 vs. 54 deaths, rate ratio 0.63, 95% CI 0.41 to 0.98) at 5-year followup, but was rated high risk of bias due to methodological shortcomings included inadequate description of randomization or allocation concealment methods, unblinded design, failure to report attrition, and failure to control for clustering affects.
- Two trials found no clear differences in mortality with US screening at 4- versus 12-month intervals, or at 3- versus 6 to month intervals.

Harms

- One study reported no serious adverse events associated with administration of gadoxetic acid for MRI and one study reported no clear differences in adverse events between CT

with contrast at 3 ml/s versus 5 ml/s. No study reported rates of adverse events associated with use of microbubble contrast agents in US, and harms were not reported in randomized trials of screening with imaging.

Detailed Synthesis

KQ1.a. What is the comparative test performance of imaging-based surveillance strategies for detecting HCC?

Ultrasound

In surveillance settings, using patients with HCC as the unit of analysis, sensitivity of US without contrast was 0.82 (95% CI 0.66 to 0.92, 3 studies) and specificity was 0.87 (95% CI 0.77 to 0.93, 2 studies), for a LR+ of 6.2 (95% CI 3.6 to 11) and LR- of 0.20 (0.10-0.40) (Appendix D; Figure 5).[51,82,85] Using HCC lesions as the unit of analysis, sensitivity was 0.60 (95% CI 0.36 to 0.80, 2 studies) and specificity 0.94 (95% CI 0.83 to 0.98, 1 study), for an LR+ of 9.8 (95% CI 3.7 to 26) and LR- of 0.43 (95% CI 0.24 to 0.74).[57,100]

In nonsurveillance settings, using patients with HCC as the unit of analysis, the sensitivity of US without contrast was 0.73 (95% CI 0.46 to 0.90, 8 studies) and specificity was 0.93 (95% CI 0.85 to 0.97, 6 studies), for a LR+ of 11 (95% CI 5.4 to 21) and LR- of 0.29 (95% CI 0.13 to 0.65) (Figures 6 and 7).[49,59,70,75,76,98,99,105] Restricting the analysis to studies that avoided a case-control design resulted in lower sensitivity (0.54, 95% CI 0.38 to 0.70, 6 studies). Other sensitivity analyses had little effect on estimates (e.g., restricted to studies conducted in United States and Europe, excluded high risk of bias studies, or restricted to studies with blinded interpretation of imaging) or resulted in imprecise estimates due to small numbers of studies (analysis restricted to prospective studies).

Using HCC lesions as the unit of analysis, the sensitivity of US without contrast was 0.60 (95% CI 0.42 to 0.75, 11 studies) (Figure 8).[49,53,65,70,71,73,76,77,89,99,105] Only two studies reported specificity, with inconsistent results (0.63, 95% CI 0.52 to 0.73[53] and 0.95, 95% CI 0.85 to 0.99[89]). For US with contrast, sensitivity was 0.76 (95% CI 0.53 to 0.90, 6 studies) (Figure 9).[48,64,66,68,73,106] Five of the contrast-enhanced studies evaluated perflubutane.[48,66,68,73,106] Sensitivity was somewhat lower when analyses were restricted to studies conducted in the United States or Europe, excluded high risk of bias studies, and avoided a case-control design, but confidence intervals were wide.

Figure 5. Test performance of ultrasound without contrast for detection of patients with hepatocellular carcinoma in surveillance settings

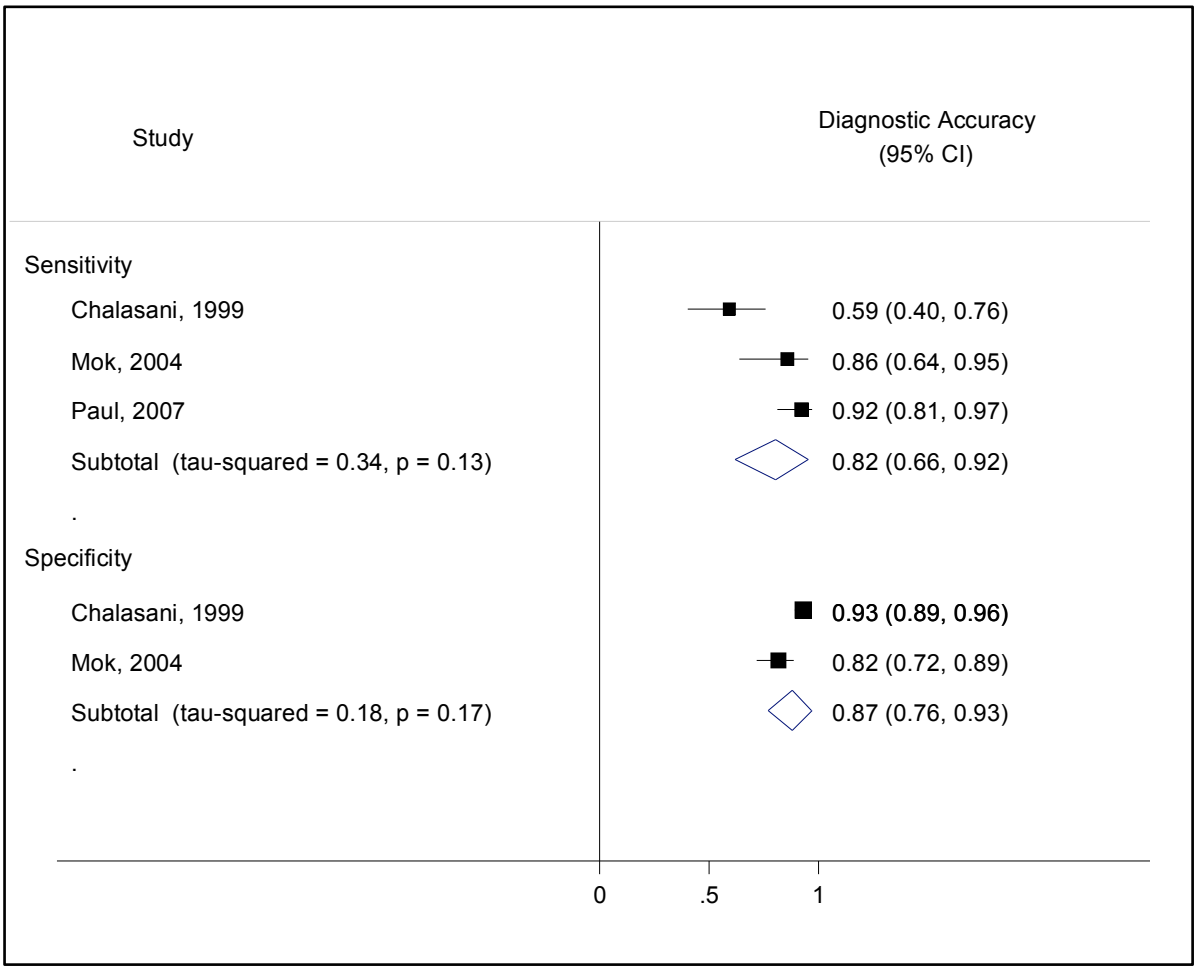


Figure 6. Sensitivity of ultrasound without contrast for detection of patients with hepatocellular carcinoma in nonsurveillance settings

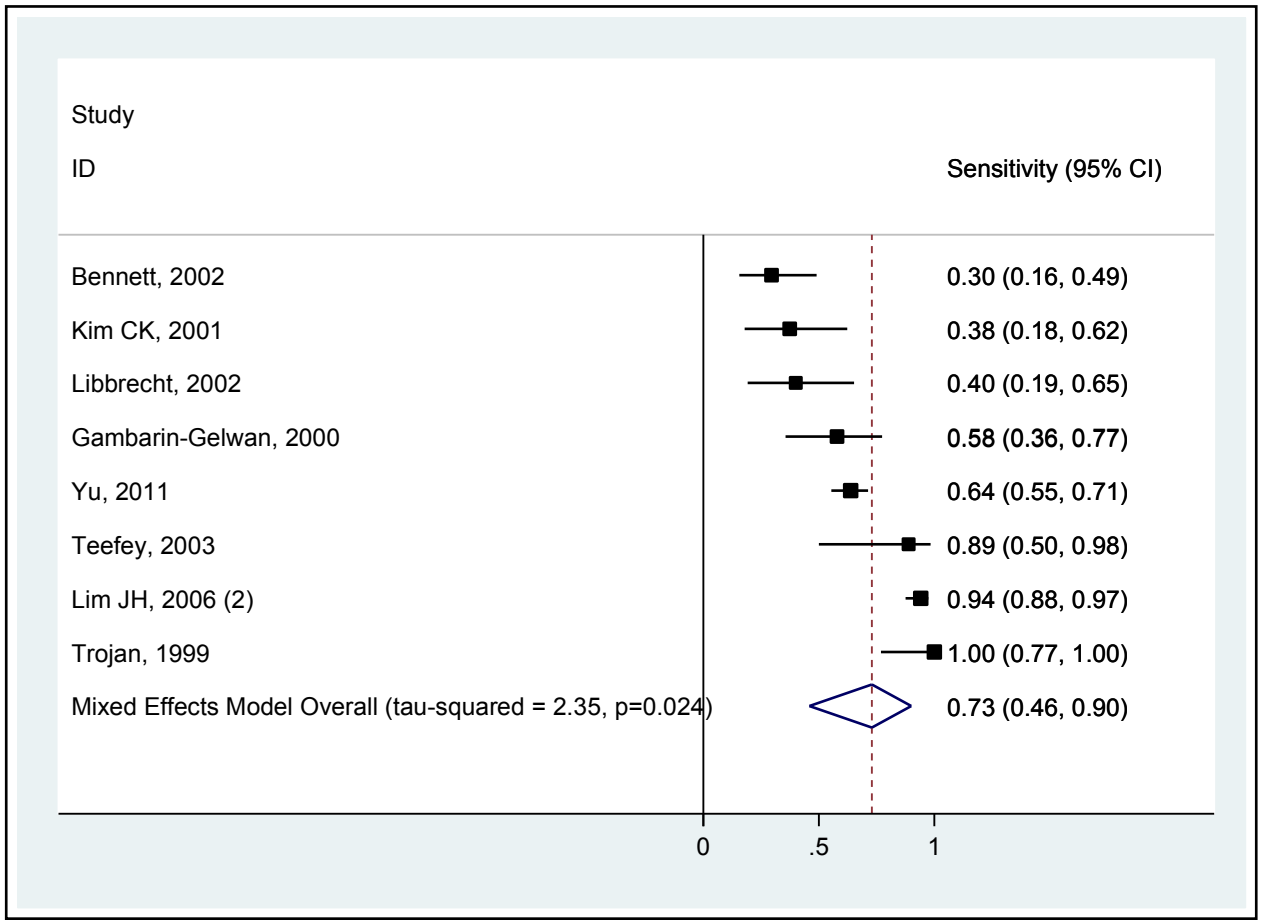


Figure 7. Specificity of ultrasound without contrast for detection of patients with hepatocellular carcinoma in nonsurveillance settings

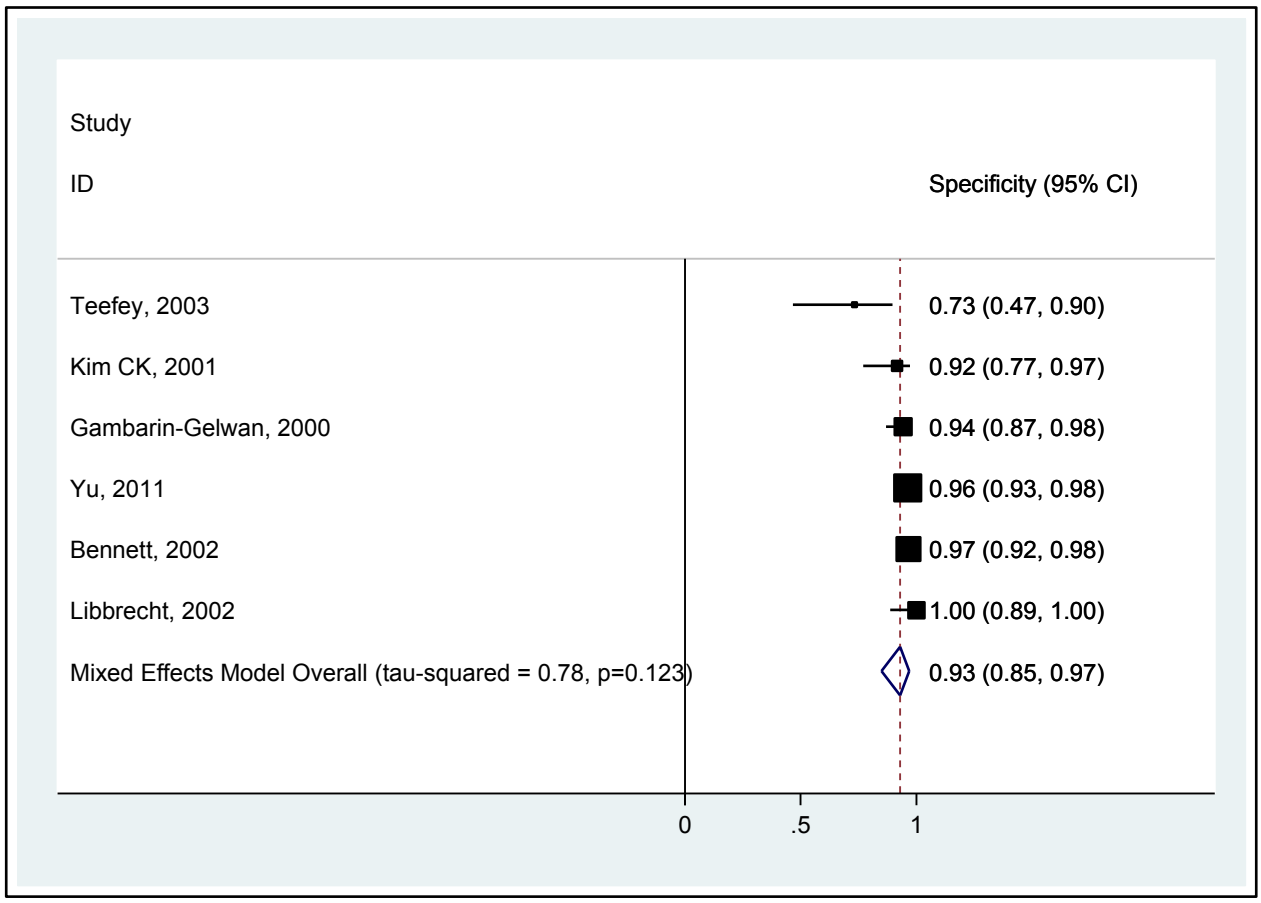


Figure 8. Sensitivity of ultrasound without contrast for detection of hepatocellular carcinoma lesions in nonsurveillance settings

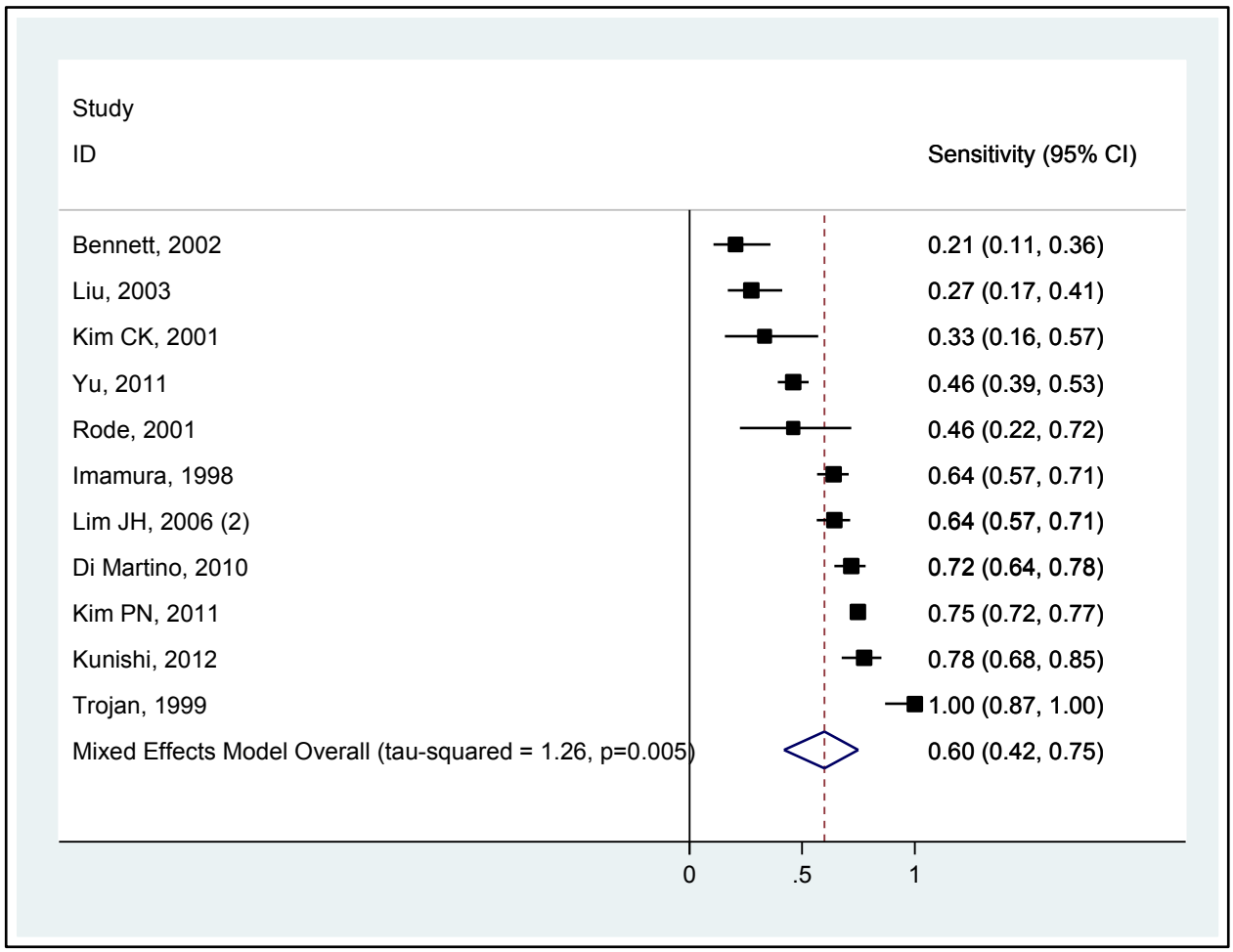
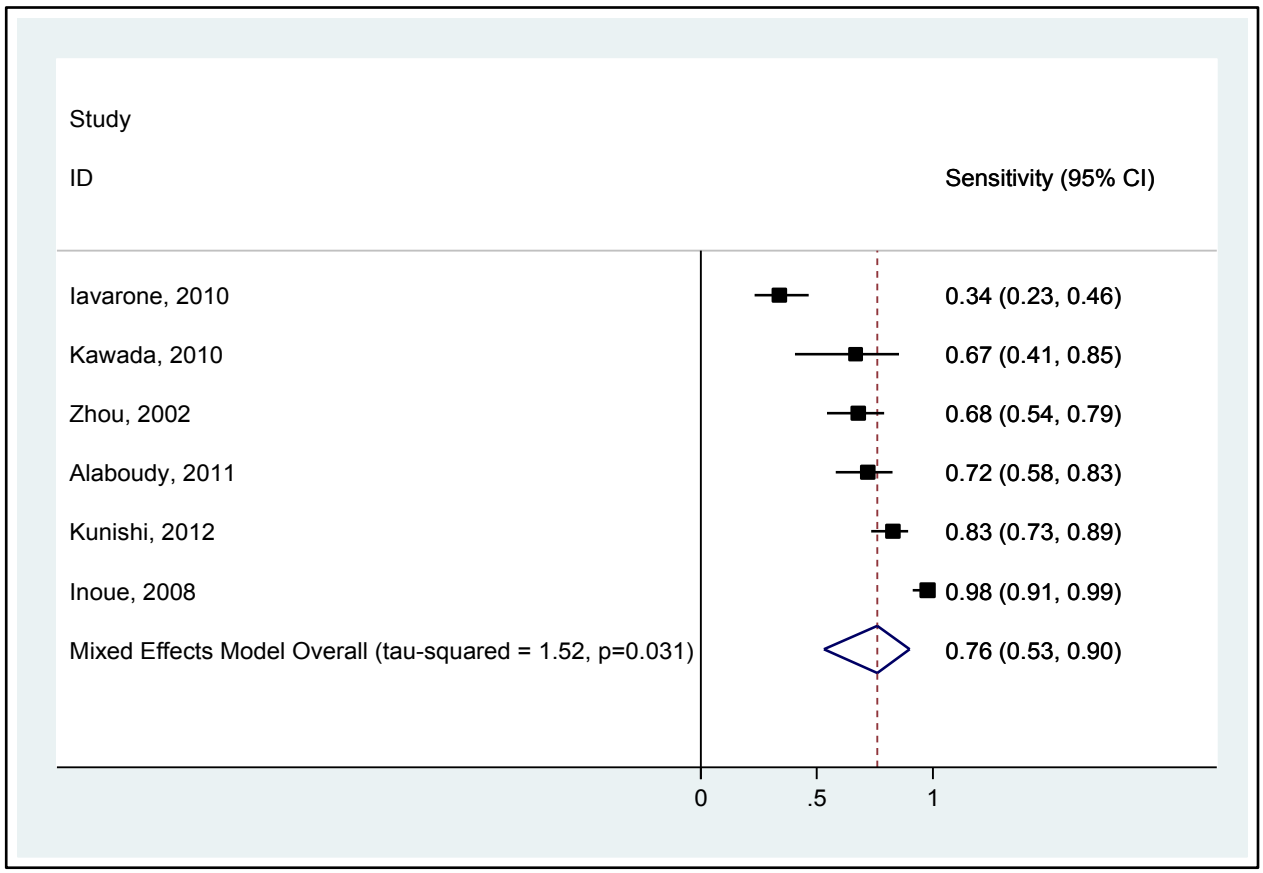


Figure 9. Sensitivity of ultrasound with contrast for detection of hepatocellular carcinoma lesions in nonsurveillance settings



Computed Tomography

Few studies evaluated CT in surveillance settings. Using patients with HCC as the unit of analysis, sensitivity was 0.84 (95% CI 0.59 to 0.95) and specificity 0.99 (95% CI 0.86 to 0.999), based on two studies (Table 7.).[51,100] In one study that used HCC lesions as the unit of analysis, sensitivity was 0.62 (95% CI 0.46 to 0.76).[57]

In nonsurveillance settings, using patients with HCC as the unit of analysis, sensitivity of CT was 0.83 (95% CI 0.75 to 0.89, 16 studies) and specificity 0.92 (95% CI 0.86 to 0.96, 11 studies), for a LR+ of 11 (95% CI 5.6 to 20) and LR- of 0.19 (95% CI 0.12 to 0.28) (Figures 10 and 11).[75,98,99,105,113,116,117,120,139,140,149,160,171,182,186,187] Using HCC lesions as the unit of analysis, sensitivity was 0.77 (95% CI 0.73 to 0.81, 75 studies) and specificity 0.89 (95% CI 0.83 to 0.93, 20 studies), for a LR+ of 7.0 (95% CI 4.6 to 11) and LR- of 0.25 (95% CI 0.21 to 0.30) (Figures 12 and 13).[41,53,64,89,105-108,110-114,116-118,120,122-128,130-135,138,140,142,143,146-154,157,160-165,167-170,173,175,177-180,183,185,187-189,192-195,197-202]. Using liver segments with HCC as the unit of analysis, sensitivity was 0.90 (95% CI 0.87 to 0.93, 7 studies) and specificity was 0.97 (95% CI 0.94 to 0.98, 7 studies), for a LR+ of 26 (95% CI 15 to 45) and LR- of 0.10 (95% CI 0.07 to 0.13).[119,136,144,145,159,190,198]

The following analyses had little impact on estimates of sensitivity and specificity or measures of heterogeneity: excluding high risk of bias studies and sensitivity analyses restricted to studies that were performed in the United States and Europe, used a prospective design, avoided a case-control design, used blinded imaging interpretation, were restricted to hypervascular HCC, or were restricted to HCC lesions <2 cm.

Magnetic Resonance Imaging

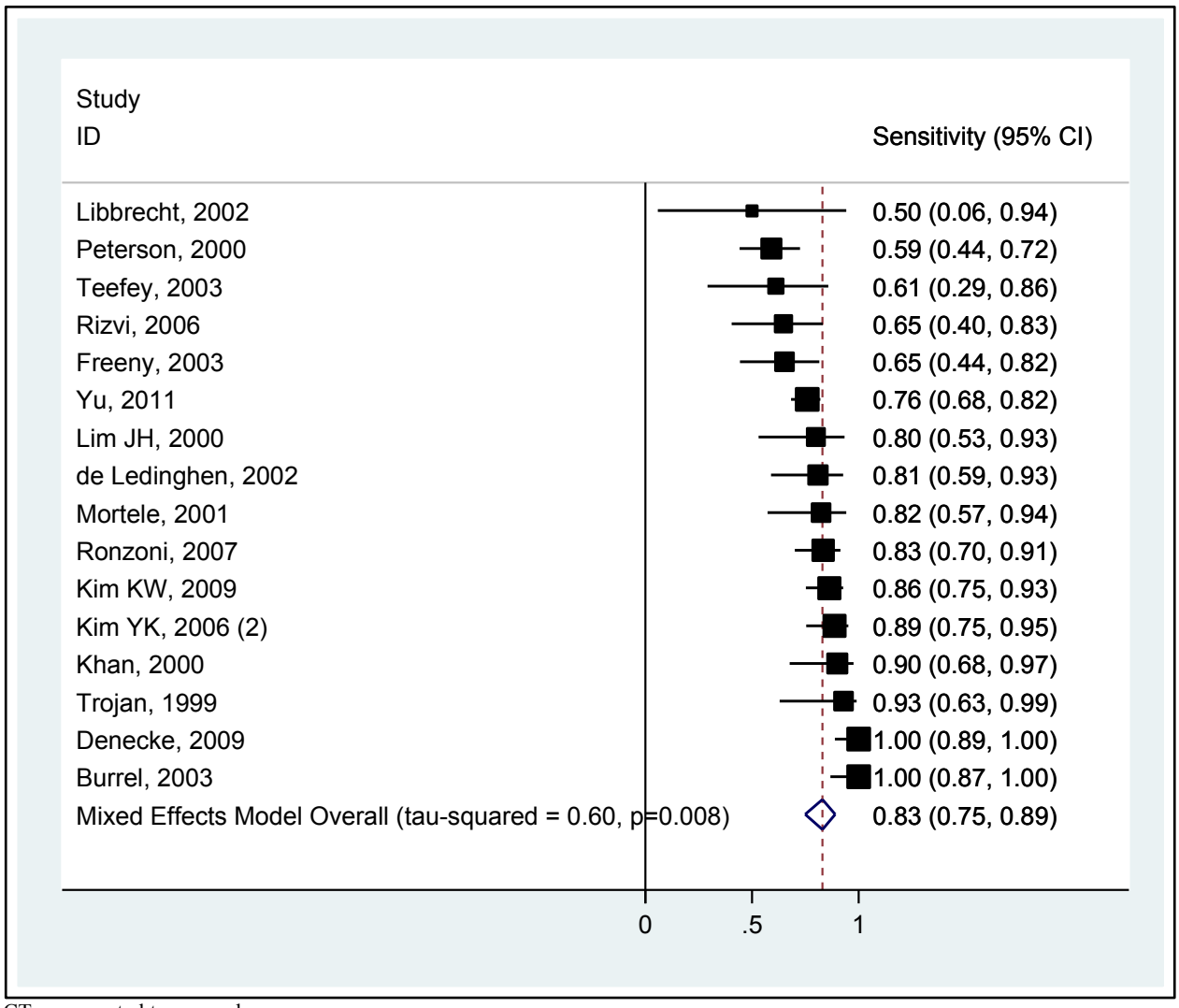
No study evaluated MRI in surveillance settings.

In nonsurveillance settings, using patients with HCC as the unit of analysis, sensitivity was 0.87 (95% CI 0.77 to 0.93, 11 studies) and specificity was 0.88 (95% CI 0.79 to 0.93, 9 studies), for a LR+ of 7.2 (95% CI 3.9 to 13) and LR- or 0.15 (95% CI 0.08 to 0.27) (Table 8; Figures 14 and 15).[75,98,105,116,237,238,241,250,253,257]

Using HCC lesions as the unit of analysis, sensitivity was 0.83 (95% CI 0.80-0.86, 68 studies) and specificity was 0.83 (95% CI 0.70-0.92, 13 studies), for a LR+ of 5.0 and LR- of 0.20 (95% CI 0.16 to 0.2.6) (Figures 16 and 17).[53,64,89,105,108,110,113,116,118,122,123,125,127,128,131,132,143,148,149,151,153,164,169,175,178,180,183,185,188,193,195,200,203,208,213,215,217,223,225,227-235,237,238,240-242,245,246,248,249,251,252,254,258,260-263,265,267,268]

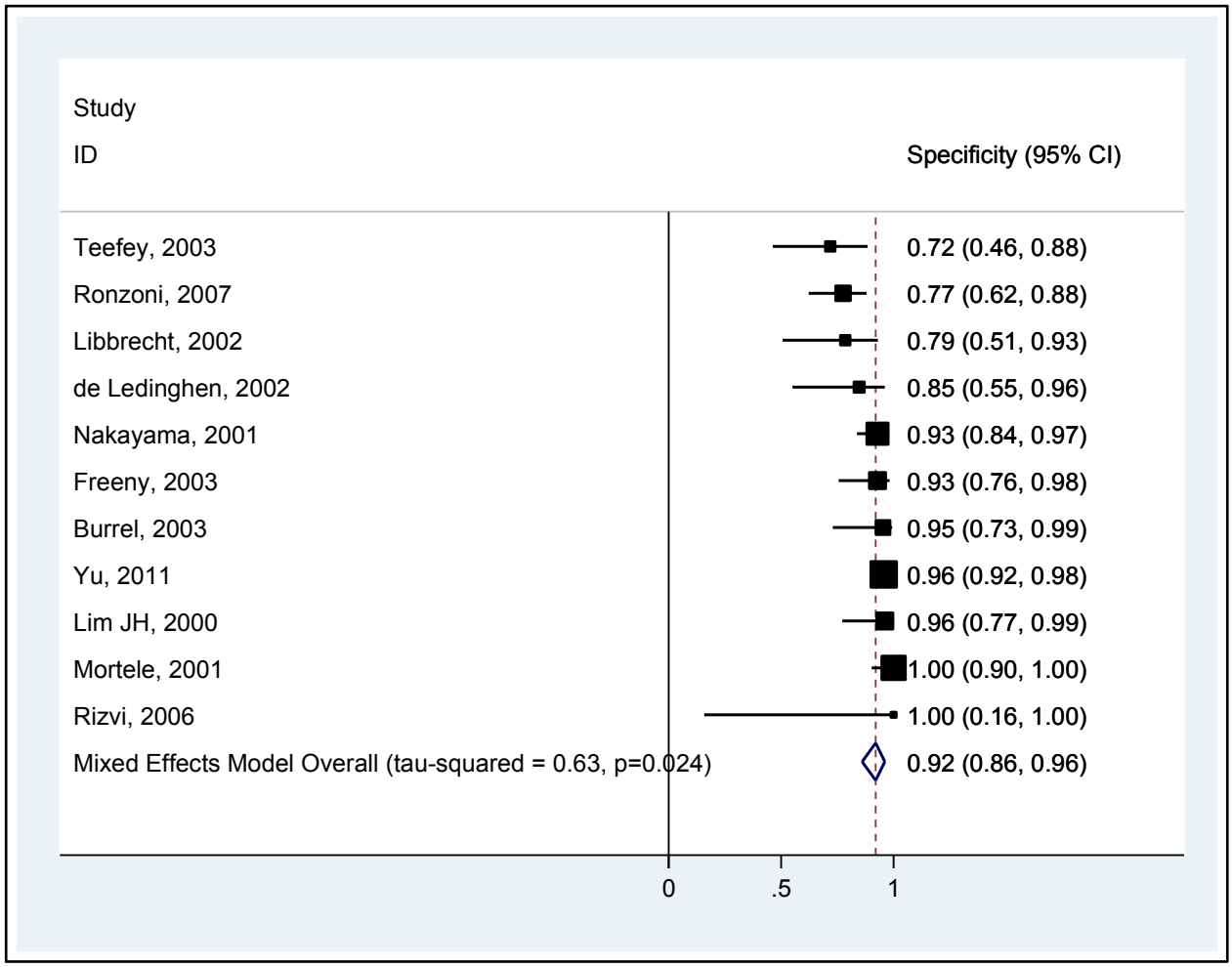
The following analyses had little impact on estimates of sensitivity and specificity or measures of heterogeneity: excluding high risk of bias studies and sensitivity analyses restricted to studies performed in the United States and Europe, that used a prospective design, avoided a case-control design, used blinding imaging interpretation, were poor quarestricted to hypervascular HCC; or studies restricted to HCC lesions <2 cm.

Figure 10. Sensitivity of CT for detection of patients with hepatocellular carcinoma in nonsurveillance settings



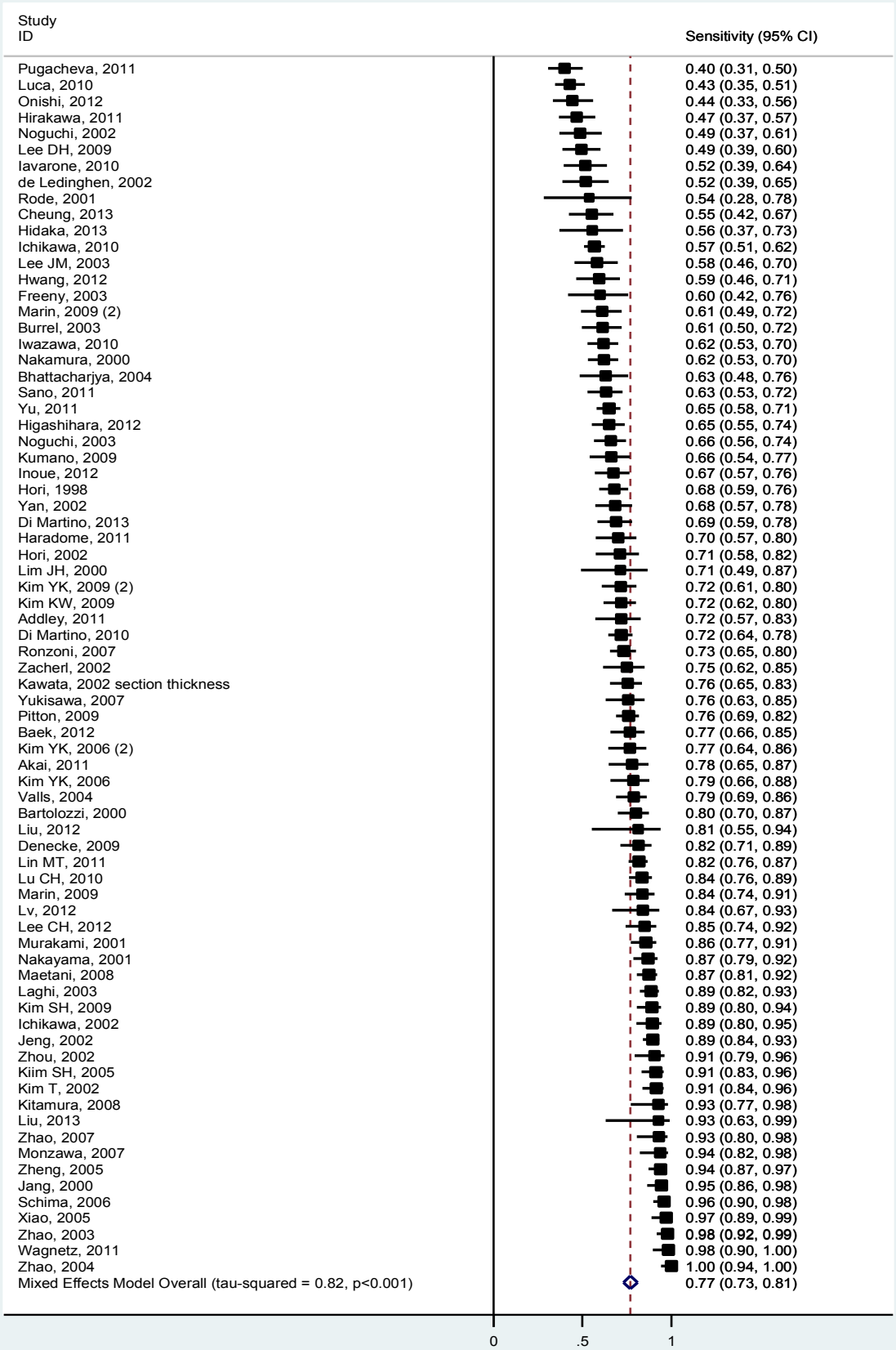
CT = computed tomography

Figure 11. Specificity of CT for detection of patients with hepatocellular carcinoma in nonsurveillance settings



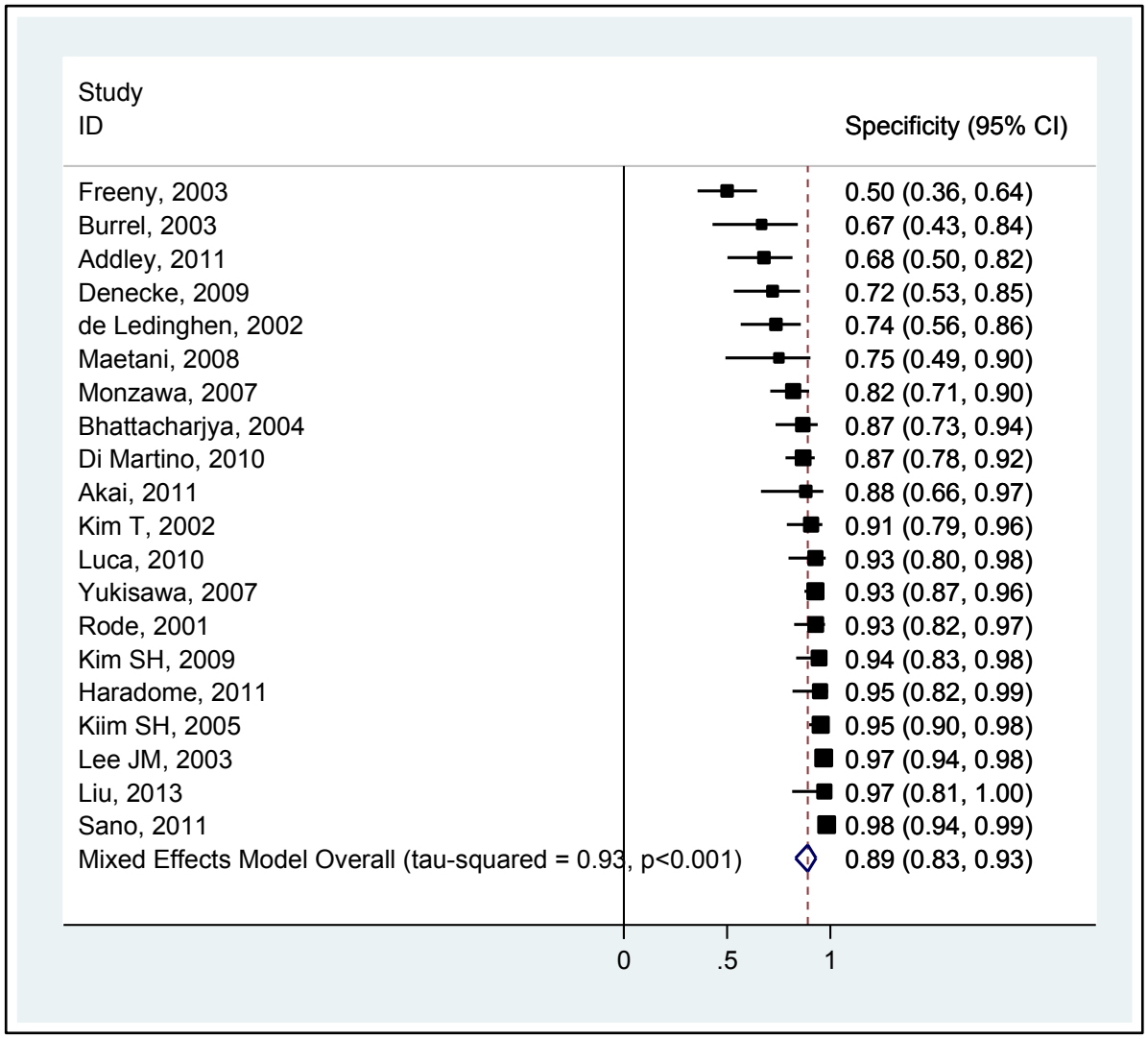
CT = computed tomography

Figure 12. Sensitivity of CT for detection of hepatocellular carcinoma lesions in nonsurveillance settings



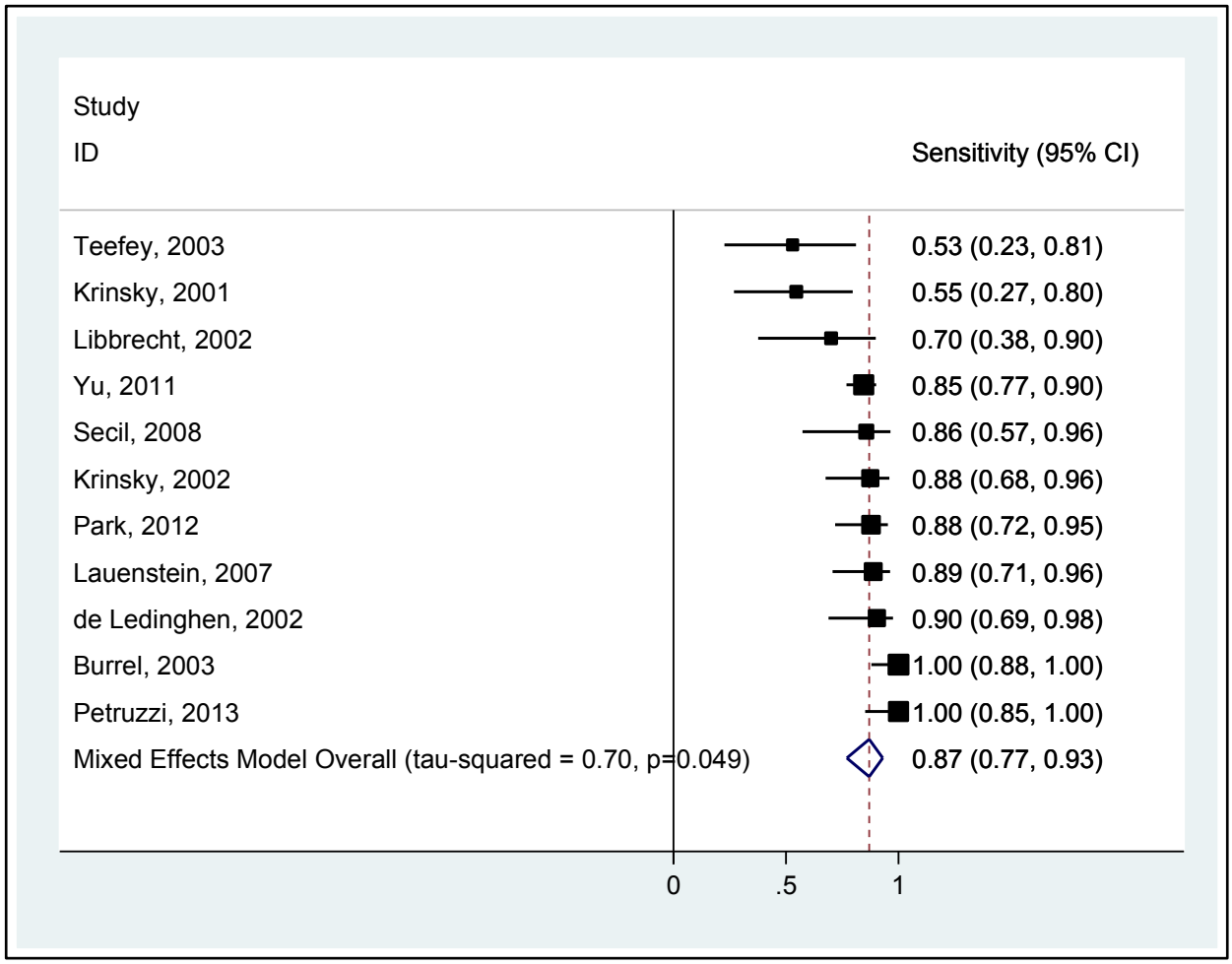
CT = computed tomography

Figure 13. Specificity of CT for detection of hepatocellular carcinoma lesions in nonsurveillance settings



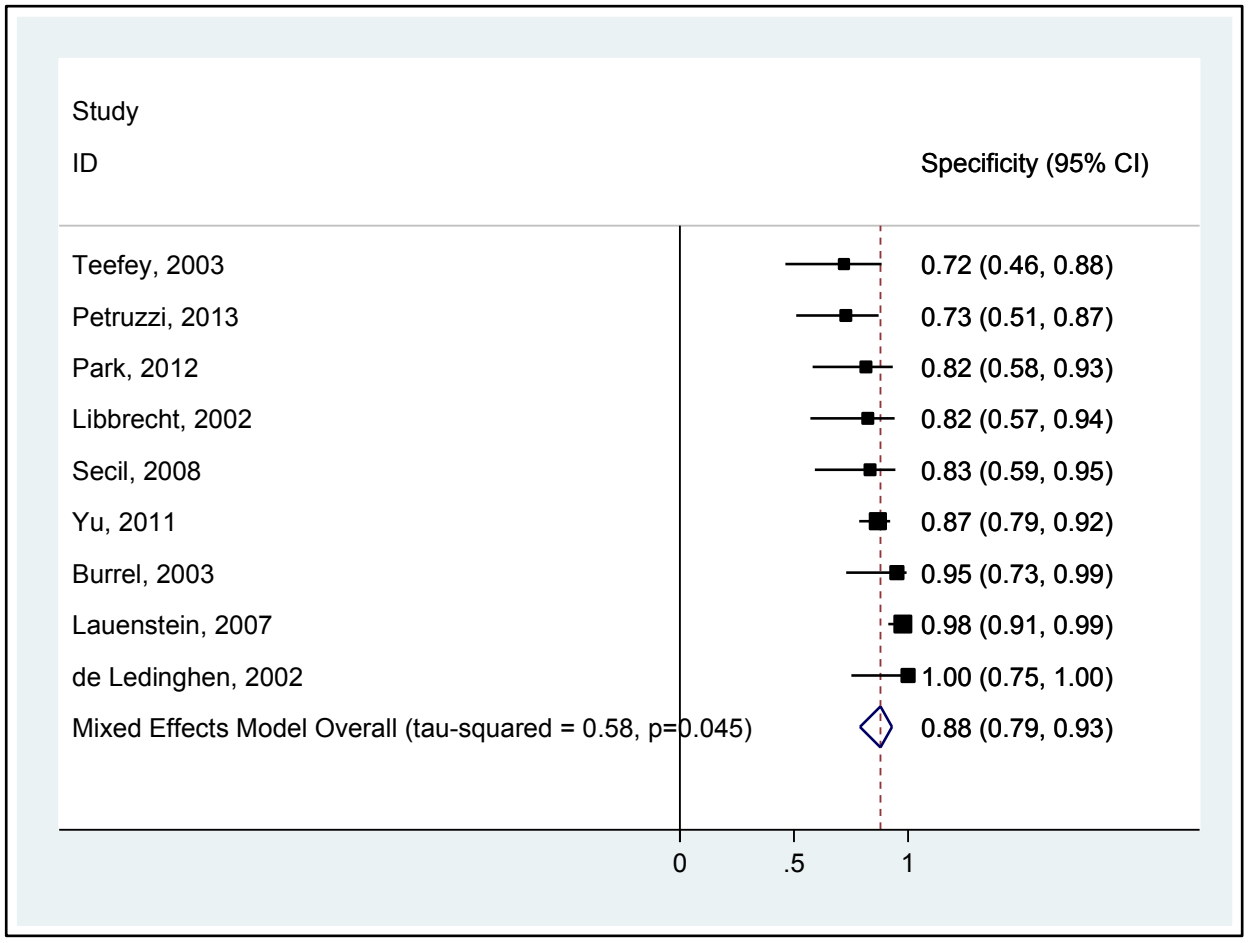
CT = computed tomography

Figure 14. Sensitivity of MRI for detection of patients with hepatocellular carcinoma in nonsurveillance settings



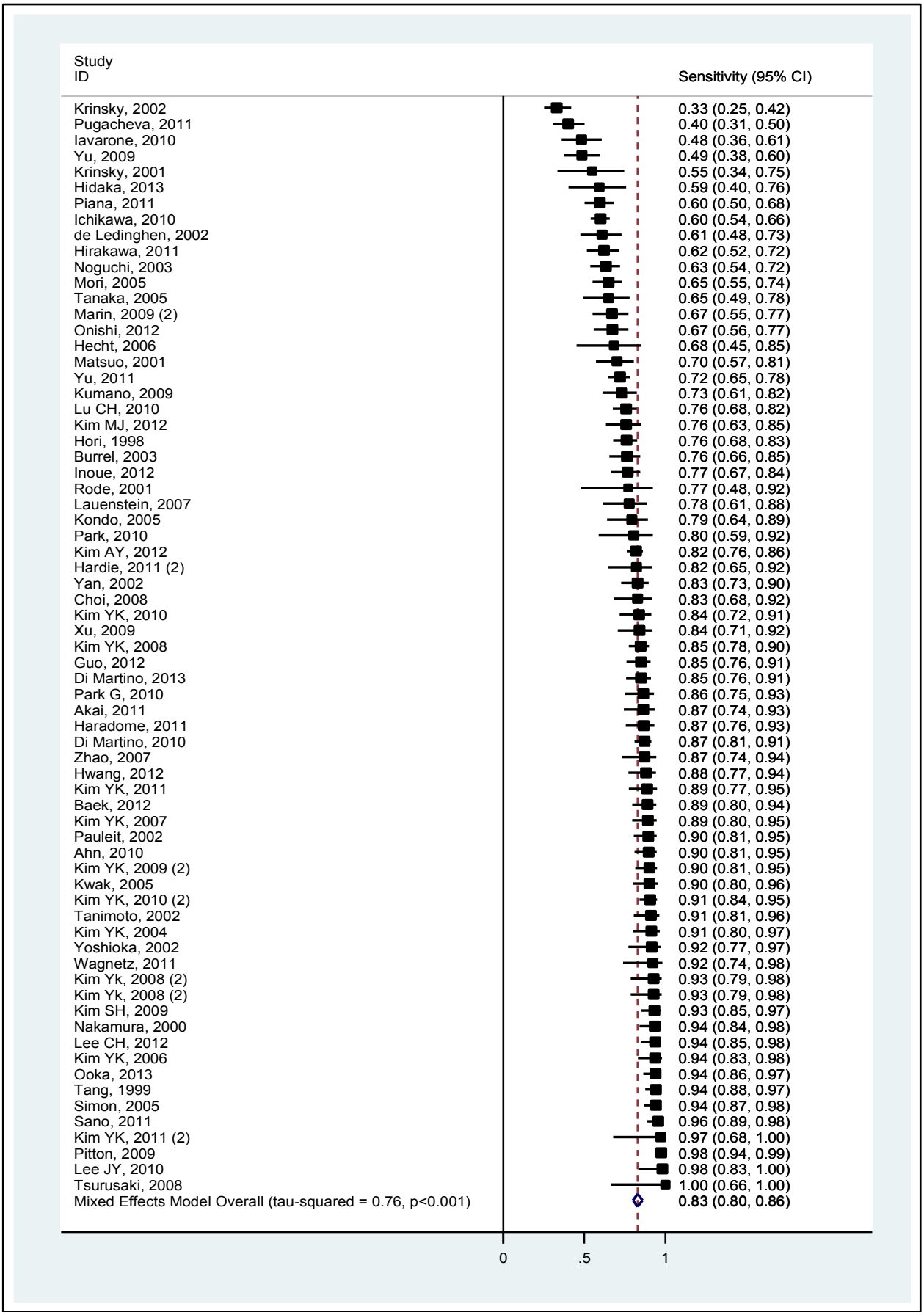
MRI = magnetic resonance imaging

Figure 15. Specificity of MRI for detection of patients with hepatocellular carcinoma in nonsurveillance settings



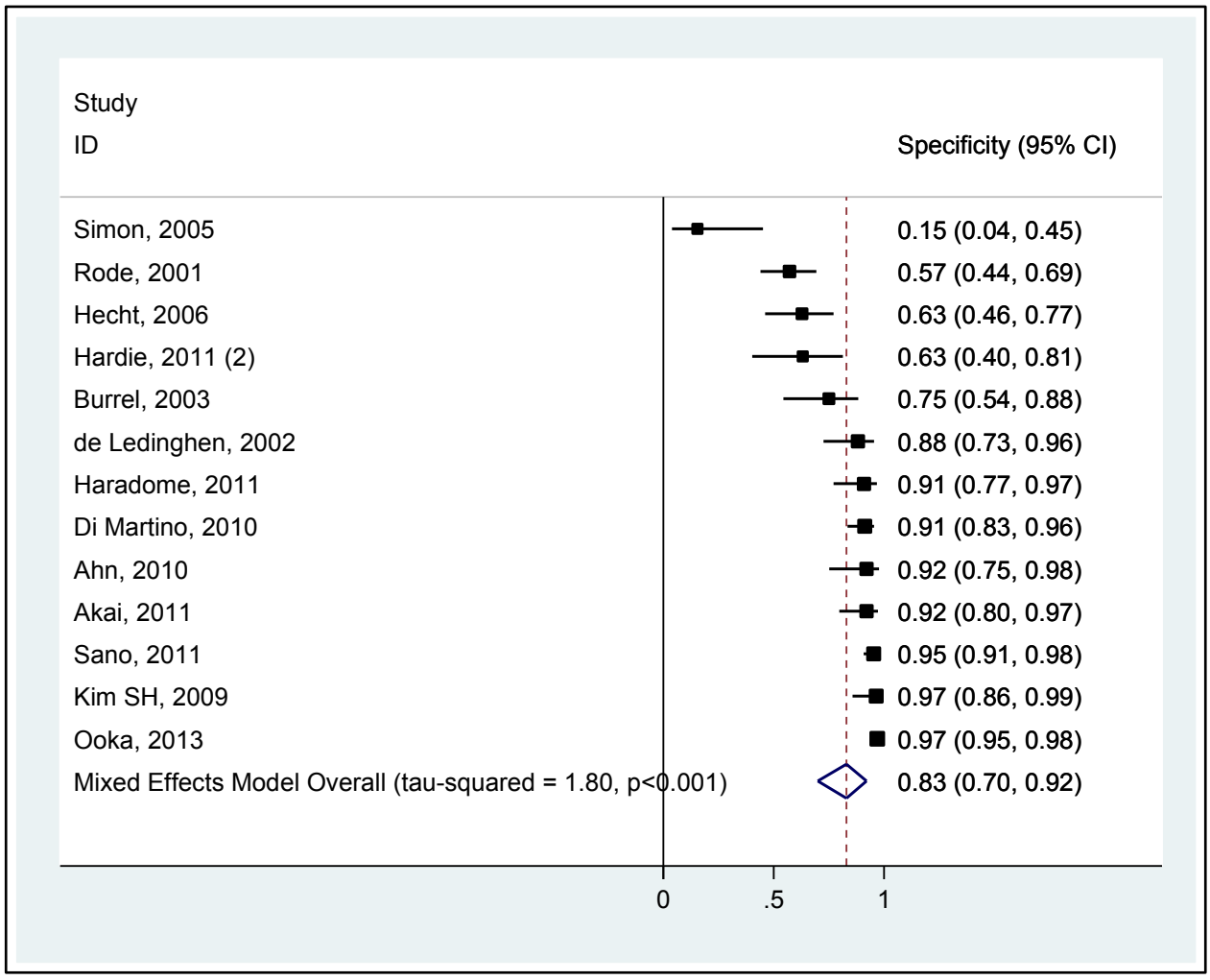
MRI = magnetic resonance imaging

Figure 16. Sensitivity of MRI for detection of hepatocellular carcinoma lesions in nonsurveillance settings



MRI = magnetic resonance imaging

Figure 17. Specificity of MRI for detection of hepatocellular carcinoma lesions in nonsurveillance settings



MRI = magnetic resonance imaging

Positron Emission Tomography

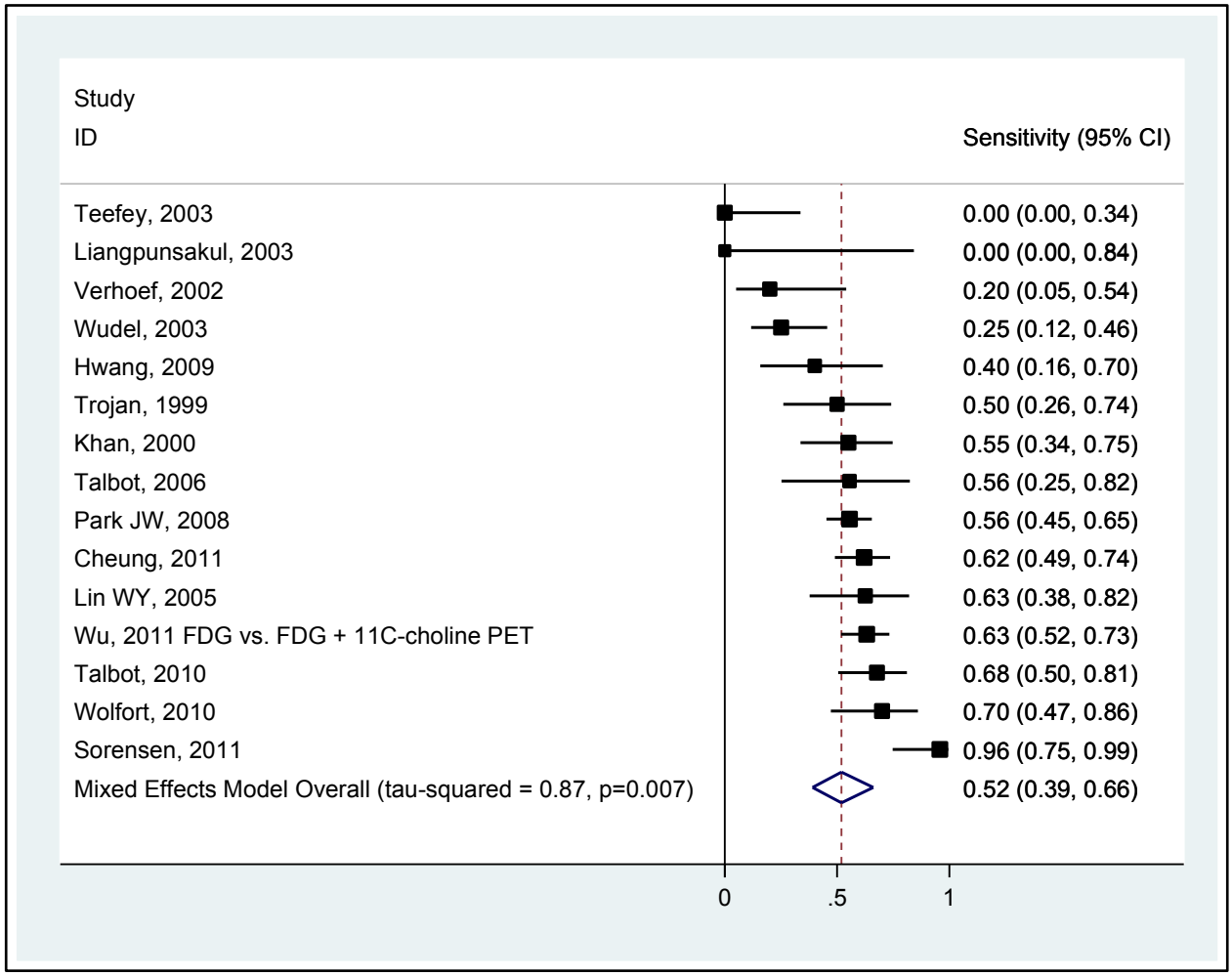
No study evaluated PET in surveillance settings.

In nonsurveillance settings, using patients with HCC as the unit of analysis, sensitivity of FDG PET was 0.52 (95% CI 0.39 to 066, 15 studies) and specificity was 0.95 (95% CI 0.92 to 0.99, 5 studies) (Appendix G; Figures 18 and 19).[98,99,139,272,277,281-284,287-292] Using HCC lesions as the unit of analysis, sensitivity was 0.56 (95% CI 0.41 to 0.69, 4 studies) and specificity 0.91 (95% CI 0.76 to 0.98, 1 study) (Figure 20).[272,283,287,293] Results were similar when analyses excluded high risk of bias studies, or when analyses were restricted to studies that used a prospective design or were conducted in the United States or Europe.

Using patients with HCC as the unit of analysis, sensitivity of ^{11}C-acetate PET was 0.85 (95% CI 0.67 to 0.94, 4 studies) (Figure 21).[272,277,280,283] Using HCC lesions as the unit of analysis, sensitivity was 0.78 (95% CI 0.61 to 0.89, 4 studies) (Figure 22).[114,280,283,293] Sensitivities of around 0.90 were reported for PET with dual tracers (FDG plus ^{11}C-acetate)[114,291] and alternative tracers such as ^{18}F-fluorothymidine[274] or ^{18}F-fluorochlorine,[287,288] but evidence was limited to one or two studies each.

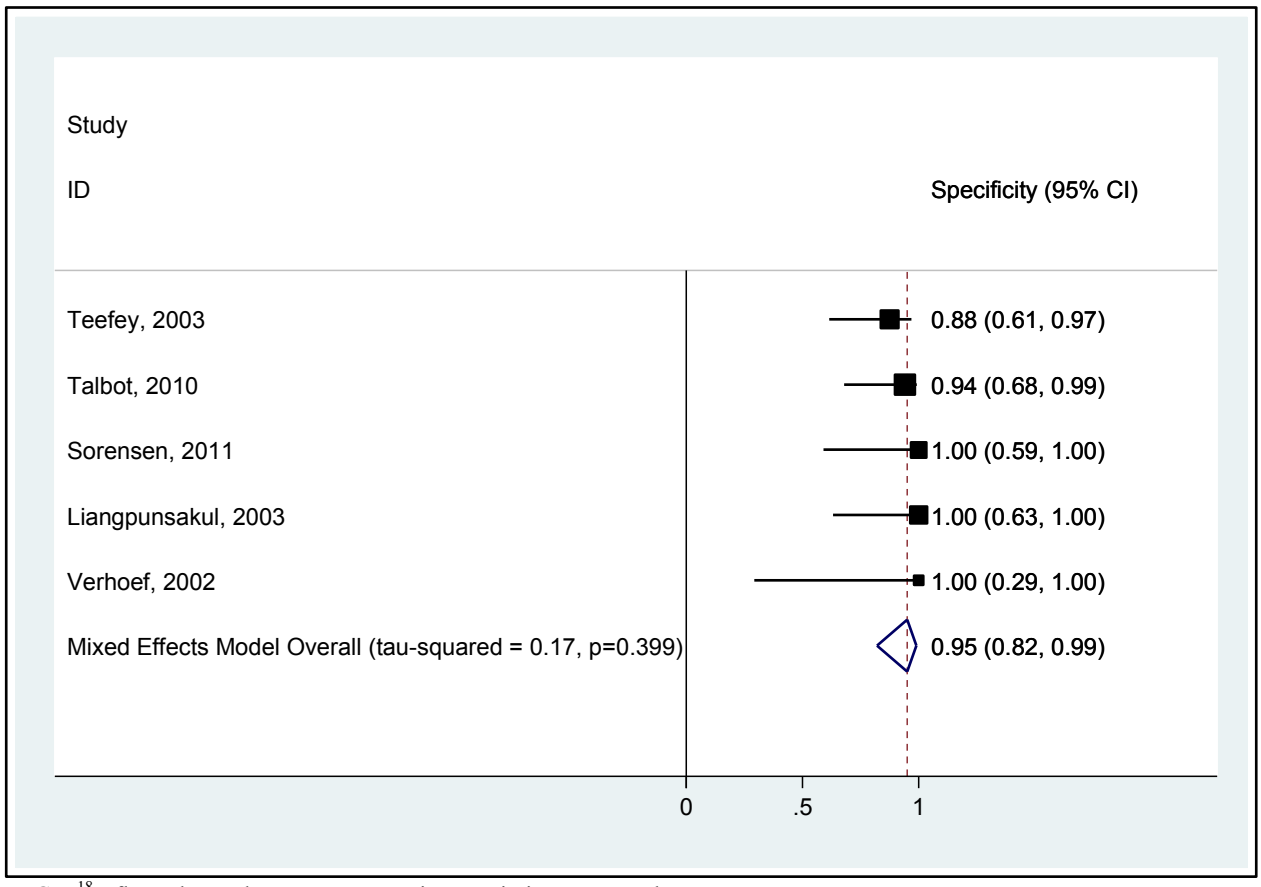
Three studies found FDG PET associated with sensitivity of 0.70 (95% CI 0.32 to 0.92, 3 studies) for detection of recurrent intrahepatic HCC, with a specificity of 0.71 (95% CI 0.29 to 0.96).[279,286,292]

Figure 18. Sensitivity of FDG PET for detection of patients with hepatocellular carcinoma in nonsurveillance settings



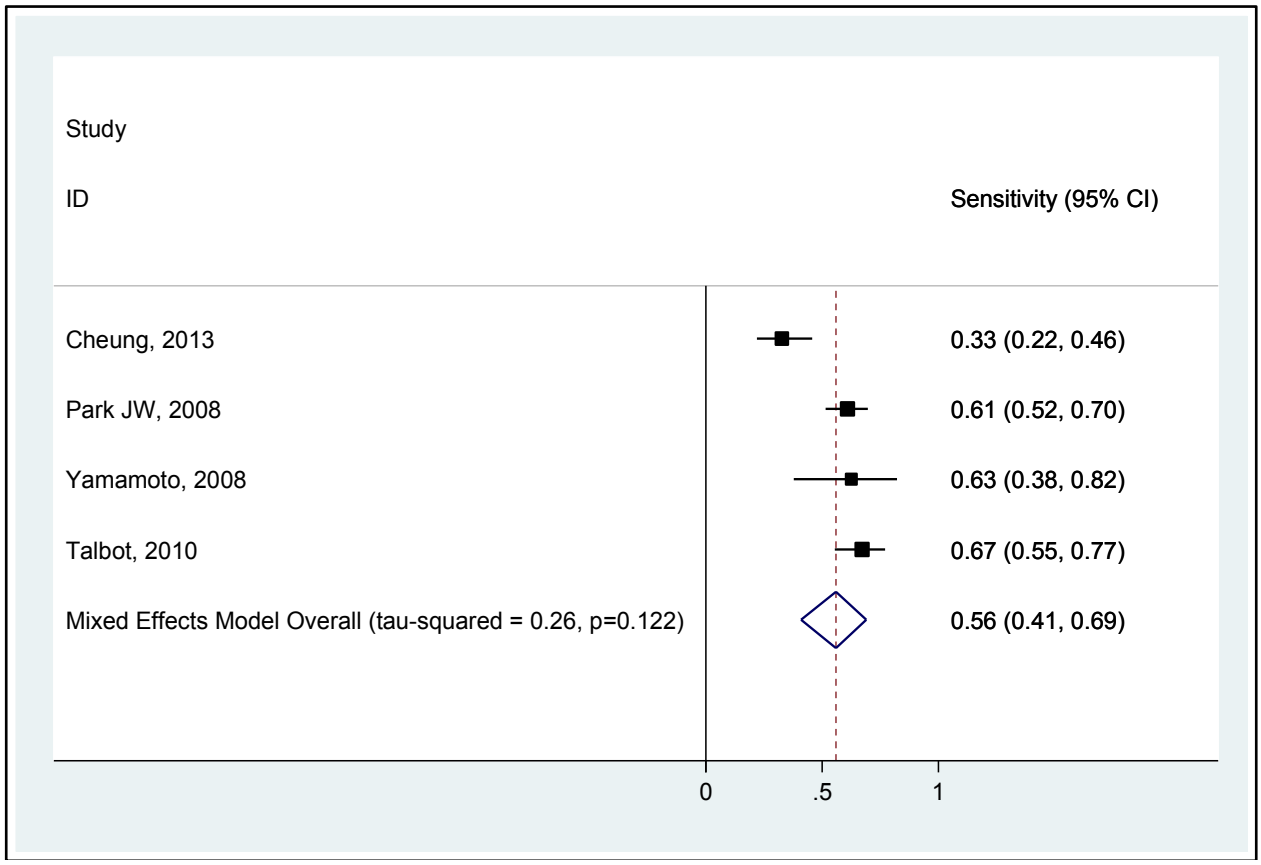
FDG = ^{18}F-fluorodeoxyglucose; PET = positron emission tomography

Figure 19. Specificity of FDG PET for detection of patients with hepatocellular carcinoma in nonsurveillance settings



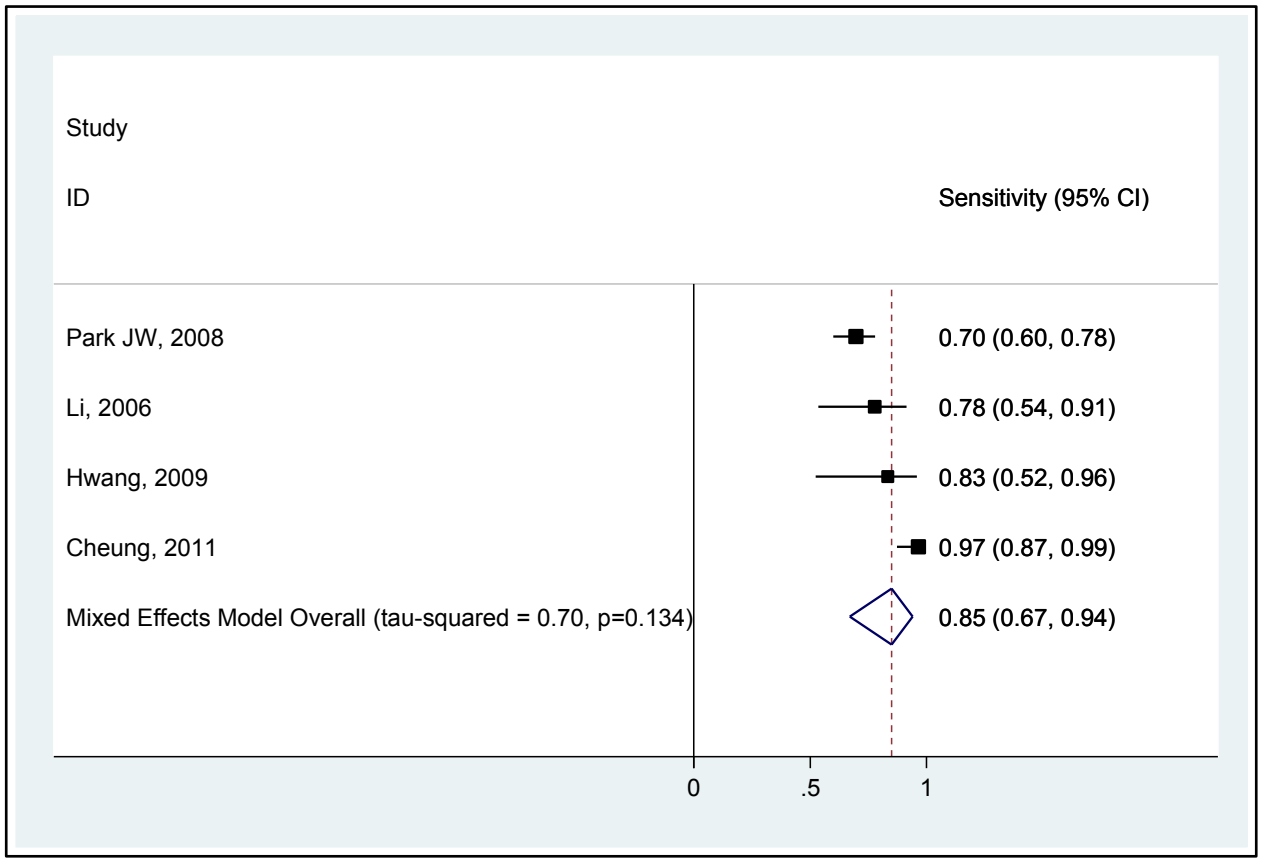
FDG = ^{18}F-fluorodeoxyglucose; PET = positron emission tomography

Figure 20. Sensitivity of FDG PET for detection of hepatocellular carcinoma lesions in nonsurveillance settings



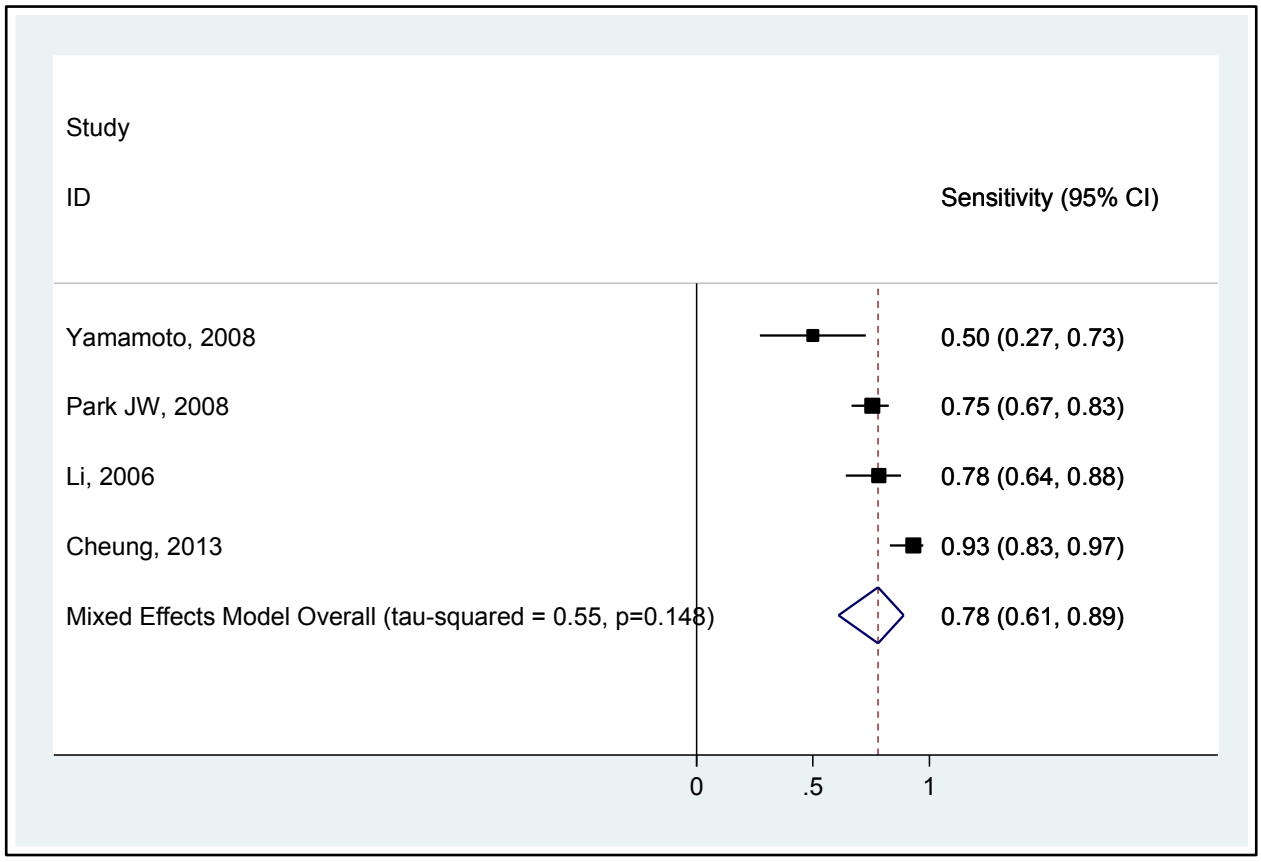
FDG = ^{18}F-fluorodeoxyglucose; PET = positron emission tomography

Figure 21. Sensitivity of ^{11}C-acetate PET for detection of patients with hepatocellular carcinoma in nonsurveillance settings



PET = positron emission tomography

Figure 22. Sensitivity of ^{11}C-acetate PET for detection of hepatocellular carcinoma lesions in nonsurveillance settings



PET = positron emission tomography

Ultrasound versus Computed Tomography

Using patients with HCC as the unit of analysis, sensitivity was lower for US without contrast (0.68, 95% CI 0.54 to 0.80) than for CT (0.80, 95% CI 0.68 to 0.88), for a difference of -0.12 (95% CI -0.20 to -0.03), based on six studies (Table 10.).[51,75,98-100,105] Findings were similar when one high risk of bias study[75] was excluded. Two of the studies were conducted in surveillance settings; both found US associated with lower sensitivity than CT (0.59 vs. 0.91[51] and 0.60 vs. 0.70[100]), with similar specificity.

Using HCC lesions as the unit of analysis, US without contrast was associated with lower sensitivity than CT (0.55, 95% CI 0.43 to 0.66 versus 0.66, 95% CI 0.54 to 0.76, for a difference of -0.11 95% CI -0.18 to -0.04), based on three studies.[53,89,105] Three studies reported similar findings for US with contrast versus CT (sensitivity 0.58, 95% CI 0.37 to 0.77 vs. 0.74, 95% CI 0.54 to 0.87, for a difference of -0.16, 95% CI -0.32 to -0.01).[48,64,106] Differences in sensitivity were similar for HCC lesions <2 cm.[53,57,64,105] None of the studies were performed in surveillance settings.

Ultrasound versus Magnetic Resonance Imaging

No study evaluated MRI versus CT in surveillance settings.

In nonsurveillance settings, using patients with HCC as the unit of analysis, three studies found US without contrast associated with lower sensitivity than MRI (0.61, 95% CI 0.48 to 0.74 vs. 0.81, 95% CI 0.69 to 0.89, for a difference of -0.19, 95% CI -0.30 to -0.08), but higher specificity (0.94, 95% CI 0.87 to 0.97 vs. 0.82, 95% CI 0.66 to 0.91, for a difference of 0.13, 95% CI 0.03 to 0.22) (Table 10).[75,98,105] Using HCC lesions as the unit of analysis, three studies found US without contrast associated with lower sensitivity than MRI (0.57, 95% CI 0.42 to 0.71 versus 0.79, 95% CI 0.67 to 0.88, for a difference of -0.22, 95% CI -0.31 to -0.14).[53,89,105] None of the studies were performed in surveillance settings.

Two studies found US with contrast associated with lower sensitivity than MRI (0.54, 95% CI 0.25 to 0.80 vs. 0.70, 95% CI 0.40-0.89, for a difference of -0.16, 95% CI -0.30 to -0.02).[48,64] There were no clear differences between US with contrast versus MRI for HCC lesions <2 cm or for well-differentiated HCC lesions.

Magnetic Resonance Imaging versus Computed Tomography

No study evaluated MRI versus CT in surveillance settings. In nonsurveillance settings, using patients with HCC as the unit of analysis, four studies found no clear differences between MRI and CT in sensitivity or specificity (Table 10).[75,98,105,113] Results were similar when high risk of bias studies were excluded.

Using HCC lesions as the unit of analysis, 28 studies found MRI associated with higher sensitivity than CT (0.81, 95% CI 0.77 to 0.84 vs. 0.72, 95% CI 0.67 to 0.77, for a difference of 0.09 (95% CI 0.06 to 0.12), with no difference in specificity.[48,53,64,89,105,108,110,113,118,122,123,127,131,132,143,148,149,151,153,161,164,169,175,178,180,183,193,195] Results were similar when high risk of bias studies were excluded. Although sensitivity for HCC lesions <2 cm was lower for both imaging modalities, the difference in sensitivity was also around 0.10.[53,64,105] Differences in sensitivity were also similar when studies were stratified according to use of nonhepatic-specific[89,105,113,127,149,151,164,178,183,193,195] or hepatic-specific contrast[48,53,64,108,110,118,122,123,131,132,143,148,153,169,180] with MRI. Specificity was lower with nonhepatic-specific MRI than CT (0.62, 95% CI 0.51 to 0.72 vs. 0.86, 95% CI 0.77 to 0.93, for a

difference of -0.24, 95% CI -0.37 to -0.11), but only two studies of nonhepatic-specific contrast reported specificity.[89,113]

Multiple Imaging Modalities

One study found sensitivity of imaging with various combinations of two imaging modalities was similar or lower than single modality imaging, based on concordant positive findings on two imaging modalities (Table 11).[48] The other study reported higher sensitivity with multiple imaging modalities than with single modality imaging, but criteria for positive results based on multiple imaging modalities were not reported.[64] Specificity was not reported in either study.

KQ1.a.i. How is a particular technique's test performance modified by use of various reference standards?

Ultrasound

There were too few studies of US in surveillance settings to evaluate effects of using different reference standards on estimate of accuracy.

In nonsurveillance settings, using patients as the unit of analysis, the sensitivity was 0.48 (95% CI 0.35 to 0.61, 5 studies) with explanted liver as the reference standard[49,59,70,75,105] and 0.95 (0.87 to 0.98, 3 studies) using a nonexplant histopathological reference standard (Table 6.).[76,98,99] Using HCC lesions as the unit of analysis, the sensitivity was 0.34 (95% CI 0.21 to 0.49, 5 studies) with explanted liver as the reference standard[49,70,77,89,105] and ranged from 0.70 to 0.85 with other reference standards (nonexplant histopathological, imaging and clinical criteria, or mixed).[48,53,64-66,68,71-73,76,99,106]

Computed Tomography

Using patients as the unit of analysis, there were no clear differences in diagnostic accuracy based on the use of different reference standards (explanted liver, other histopathological reference standard, or mixed histological and clinical/imaging), with sensitivity ranging from 0.81 to 0.88 (Table 7).[75,98,99,105,113,116,117,120,139,140,149,160,171,177,182,186,187] Using HCC lesions as the unit of analysis, studies using explanted livers as the reference standard reported a lower sensitivity (0.69, 95% CI 0.60-0.77, 21 studies)[41,89,107,108,112,113,116,117,120,123,125,160,162-164,167,187,192,198] than studies that used a nonexplant histopathological reference standard (0.85, 95% CI 0.77 to 0.90, 13 studies)[64,106,114,122,130,132,134,143,144,161,170,188,193] or studies that used a mixed histological and clinical/imaging reference standard (0.79, 95% CI 0.74 to 0.84, 38 studies).[53,110,111,118,126,128,131,133,135,138,140,146-153,157,165,168,169,173,177-180,183,185,189,194,195,197,199-202]
Estimates of specificity stratified by the reference standard were imprecise due to small numbers of studies.

Magnetic Resonance Imaging

There were too few studies with patients as the unit of analysis that used a nonexplant reference standard to evaluate effects of different reference standards. Nine of 11 studies used an explanted liver reference standard, with a pooled sensitivity of 0.89 (95% CI 0.79 to 0.94).[75,105,113,116,237,238,241,250,253]

Using HCC lesions as the unit of analysis, studies using explanted livers as the reference standard reported a lower sensitivity (0.69, 95% CI 0.59 to 0.77, 15 studies)[89,105,108,113,116,123,125,164,208,215,217,237,238,241,268] than studies that used a nonexplant histopathological reference standard, clinical/imaging reference standard, or mixed histological

and imaging/clinical reference standard (sensitivity estimates ranged from 0.85 to 0.88) (Table 8).[53,64,110,118,122,127,128,131,132,143,148,149,151,153,169,175,178,180,183,185,188,193,195,200,203,213,223,225,227-235,240,242,245,246,248,249,251,252,254,258,260-263,265,267] Estimates of specificity stratified by the reference standard were imprecise due to small numbers of studies.

Positron Emission Tomography

No study of FDG PET used explanted livers as the reference standard. Using patients as the unit of analysis, there were no clear differences in sensitivity between studies that used a nonexplant histological reference standard (0.46, 95% CI 0.28 to 0.65, 7 studies)[98,99,139,272,283,289,290] and studies that used a mixed histological and imaging/clinical criteria reference standard (0.58, 95% CI 0.40-0.75, 8 studies), based on relatively wide and overlapping confidence intervals (Table 9). Three of the four studies that used HCC lesions as the unit of analysis used a nonexplant histological reference standard; the pooled sensitivity from this subset of studies was similar to the overall pooled estimate.[272,283,293]

KQ1.a.ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?

Ultrasound

In two studies that directly compared US with versus without contrast, there was no clear difference in sensitivity (-0.04, 95% CI -0.11 to 0.04) (Table 12).[62,73] Excluding studies that used Doppler had little effect on estimates of sensitivity and specificity, and one study[57] that directly compared use of Doppler versus no Doppler showed no clear effect on estimates of sensitivity.

In 14 studies that reported accuracy of US stratified by HCC lesion size, sensitivity was greater for lesions >2 cm (0.88, 95% CI 0.78 to 0.94) than for lesions <2 cm (0.49, 95% CI 0.31 to 0.67), for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51).[49,53,58,61,62,64,65,70,74,76,77,102,104,105] Differences were larger in studies of noncontrast US (0.48, 95% CI 0.39 to 0.51, 9 studies)[49,53,62,65,70,74,76,77,105] than in studies of US with contrast (0.17, 95% CI 0.03 to 0.32, 5 studies)[58,61,64,102,104] but these findings are difficult to interpret because sensitivity for HCC lesions <20 mm was much lower in the studies of noncontrast US (0.34, 95% CI 0.19 to 0.53) than in studies of US with contrast (0.77, 95% CI 0l.53 to 0.91). For US without contrast, sensitivity progressively improved from 0.09 (95% CI 0.02 to 0.29, 4 studies) for lesions <10 mm to 0.50 (95% CI 0.23 to 0.78, 4 studies) for lesions 10-20 mm and 0.88 (95% CI 0.66 to 0.96, 4 studies) for lesions >20 mm , for a difference of 0.37 (95% CI 0.18 to 0.57) for lesions >20 mm vs. 10-20 mm, and 0.41 (95% CI 0.19 to 0.63) for lesions 10-20 mm vs. <10 mm (Table 13).[49,53,74,76] For US with contrast, three studies found sensitivity lower for lesions 10-20 mm (0.64, 95% CI 0.33 to 0.87) than >20 mm (0.91, 95% CI 0.71 to 0.98), for a difference of 0.26 (95% CI 0.04 to 0.48).[58,64,102]

In three studies, sensitivity was 0.83 (95% CI 0.55 to 0.95) for moderately- or poorly-differentiated HCC lesions versus 0.43 (95% CI 0.15 to 0.76) for well-differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.17 to 0.64).[50,64,81] Lesion depth and body mass index had no effect on estimates of sensitivity (Table 6). Two studies reported conflicting results for effects of cirrhosis on estimates of sensitivity, with one study reporting presence of cirrhosis associated with lower sensitivity than in patients without cirrhosis[71] but the other with slightly higher sensitivity.[76] Evidence on effects of liver volume, subcapsular location, presence

of ascites, and underlying condition on estimates of accuracy was very sparse and showed no clear differences.[71,77]

Computed Tomography

Using patients with HCC as the unit of analysis, studies with a contrast rate ≥3 ml/s reported a higher sensitivity (0.87, 95% CI 0.77 to 0.93, 8 studies)[98,113,116,117,120,149,160,177,187] than studies with a contrast rate <3 ml/s (0.71, 95% CI 0.50-0.85, 4 studies),[75,105,171,182] with similar specificity, but there was no clear difference in studies that used HCC lesions as the unit of analysis (0.79, 95% CI 0.75 to 0.83, 62 studies and 0.72, 95% CI 0.58 to 0.83, 8 studies, respectively) (Table 14).

Using patients with HCC as the unit of analysis, studies with delayed phase imaging reported somewhat higher sensitivity (0.89, 95% CI 0.81 to 0.94, 7 studies)[75,99,105,116,139,171,182] than studies without delayed phase imaging (0.78, 95% CI 0.66 to 0.87, 7 studies),[98,113,117,140,149,160,187] but there was no clear difference in studies that used HCC lesions as the unit of analysis (0.75, 95% CI 0.69 to 0.80, 44 studies and 0.81, 95% CI 0.75 to 0.86, 27 studies, respectively) (Table 7).

The type of CT scanner (≥8-row multidetector, <8-row multidetector, or nonmultidetector) had no clear effect on estimates of diagnostic accuracy. Based on three studies that directly compared spectral versus standard CT, there was no clear difference in estimates of diagnostic accuracy.[165,166,196] Two studies compared effects of quantitative versus qualitative methods for evaluation of CT imaging findings on estimates of diagnostic accuracy.[140,163] In one study, use of quantitative arterial enhancement fraction mapping was associated with higher sensitivity than qualitative assessment for all HCC lesions, as well as lesions ≤2 cm.[140] In the other study, use of the percentage attention ratio threshold had no clear effect on sensitivity.[163]

In 33 studies that reported accuracy of CT stratified by HCC lesion size, sensitivity was greater for lesions >2 cm (0.94, 95% CI 0.91 to 0.95) than for lesions <2 cm (0.62, 95% CI 0.56 to 0.68), for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36).[41,53,61,104,105,107,110,112,113,117,118,125,127,129,132,133,140,142-145,152,154,159,168,169,173,177-180,187,192,197] Estimates were similar when the analysis was restricted to seven studies that met minimum technical criteria.[53,118,143,168,169,180,197] Sensitivity progressively improved from 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10-20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 versus 10-20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10-20 versus <10 mm (Table 13).[41,53,64,110,112,113,117,121,125,127,128,133,142-145,159,161,177,180,185,187,197]

In five studies that reported accuracy of CT stratified by degree of tumor differentiation, sensitivity was greater for moderately- or poorly-differentiated HCC lesions (0.82, 95% CI 0.66 to 0.91) than for well-differentiated lesions (0.50, 95% CI 0.29 to 0.70), for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45).[50,57,64,154,177]

Magnetic Resonance Imaging

There were no clear differences in estimates of diagnostic accuracy when studies were stratified according to MRI scanner type (1.5 vs. 3.0 T), type of contrast (gadopentetate or gadodiamide vs. gadoxetic acid or gadobenate), imaging phases evaluated (with or without delayed phase imaging), or timing of delayed phase imaging (>120 seconds vs. <120 seconds) (Table 8). Sensitivity was somewhat higher in studies with enhanced section thickness of ≤5 mm than in studies with section thickness >5 mm, but confidence intervals were wide and overlapped. Relatively few studies evaluated 3.0 T MRI[110,128,132,143,153,180,203,213,223,225,242,249] or MRI without delayed phase imaging,[116,151,175,263] precluding strong conclusions.

In studies that directly compared diagnostic accuracy of MRI for HCC lesions using different types of contrast, hepatic specific contrast agents (gadoxetic acid or gadobenate) were associated with slightly higher sensitivity than nonhepatic-specific contrast agents (gadopentetate or gadodiamide) (0.82, 95% CI 0.71 to 0.90 vs. 0.75, 95% CI 0.61 to 0.85, difference 0.07, 95% CI 0.01 to 0.14, 5 studies), with no difference in specificity (Table 15).[110,122,169,180,203] In studies restricted to HCC lesions <2 cm in diameter, the difference was somewhat larger (sensitivity 0.77, 95% CI 0.68 to 0.84 vs. 0.62, 95% CI 0.52 to 0.71, difference 0.15, 95% CI 0.08 to 0.22, 7 studies).[53,110,122,169,180,203,220]

In studies that directly compared diagnostic accuracy of MRI with versus without diffusion-weighted imaging, there was no difference in sensitivity.[214,223,227,247,250,254,264,266] Restricted to HCC lesions <2 cm in diameter, diffusion-weighted imaging was associated with slightly higher sensitivity (0.78, 95% CI 0.62 to 0.88 vs. 0.67, 95% CI 0.50-0.81, difference 0.10, 95% CI 0.02 to 0.18, 5 studies).[224,225,250,254,264]

In 25 studies that reported accuracy of MRI stratified by HCC lesion size, sensitivity was greater for lesions >2 cm (0.96, 95% CI 0.93 to 0.97) than for lesions <2 cm (0.65, 95% CI 0.57 to 0.73), for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37)[53,105,110,113,118,121,125,127,132,143,169,178,180,207,208,213,237,238,241,244,248,250,254,261]. The difference was greater in studies of nonhepatic-specific contrast (0.40, 95% CI 0.30-0.49, 15 studies) than in studies of hepatic-specific contrast (0.20, 95% CI 0.13 to 0.27, 9 studies). Sensitivity progressively improved from 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10-20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 versus 10-20 mm and 0.34 (95% CI 0.27 to 0.41) for 10-20 versus <10 mm (Table 13).[53,64,110,113,121,125,127,128,143,180,185,203,207,208,225,237,238,250,268]

In two studies that reported accuracy of MRI stratified by degree of tumor differentiation, sensitivity was greater for moderately- or poorly-differentiated HCC lesions (0.54, 95% CI 0.26 to 0.79) than for well-differentiated lesions (0.38, 95% CI 0.17 to 0.64), but the difference was not statistically significant (0.16, 95% CI -0.11 to 0.43).[64,123] In two studies, sensitivity decreased as Child-Pugh class increased (class A 0.97, 95% CI 0.90-0.99, class B 0.91, 95% CI 0.74 to 0.97, class C 0.79 (0.54 to 0.93).[223,261]

Positron Emission Tomography

In studies that directly compared accuracy of PET using different tracers, FDG PET was associated with lower sensitivity that ^{11}C-acetate PET when either patients (0.58 vs. 0.81, for a difference of -0.23, 95% CI -0.34 to -0.13, 3 studies[272,277,283]) or HCC lesions (0.52 vs. 0.79, for a difference of -0.27, 95 %CI -0.36 to -0.17, 3 studies[272,283,293]) were the unit of analysis. FDG PET was also associated with lower sensitivity than dual tracer PET with FDG and ^{11}C-acetate[114,291] or ^{18}F-choline PET,[287,288] but evidence was limited to 1 or 2 studies for each of these comparisons.

Using patients as the unit of analysis, sensitivity of FDG PET (0.39, 95% CI 0.24 to 0.56, 8 studies)[98,99,139,281,282,289,290,292] was lower than sensitivity of FDG PET/CT (0.65, 95% CI 0.50-0.78, 7 studies) (Table 16).[272,277,283,284,287,288,291]. Similar findings were seen in studies that used ^{11}C-acetate as the tracer and HCC lesions as the unit of analysis, but the number of studies was small (0.68, 95% CI 0.46 to 0.84, 2 studies[280,293] versus 0.85, 95% CI 0.67 to 0.94, 2 studies[114,283]).

In five studies that reported accuracy of FDG PET stratified by HCC lesion size, sensitivity was consistently higher for larger lesions (Table 17).[99,272,279,283,290] Data were not pooled due to

differences in the tumor size categories evaluated, with small samples in some studies. One study reported a similar pattern for ^{11}C-acetate PET, though the difference was less pronounced, due to higher sensitivity for lesions 2 to 5 cm in diameter.[283] Another study reported high sensitivity of ^{11}C-acetate PET for lesions ≤5 cm or >5 cm.[272]

Five studies of FDG PET found lower sensitivity for more poorly-differentiated lesions than for more well-differentiated lesions (0.49, 95% CI 0.41 to 0.58 vs. 0.78, 95% CI 0.70-0.85, for a difference of -0.29, 95% CI -0.58 to -0.41) (Table 18. PET Tumor differentiation).[99,283,287,290,291] In two studies of ^{11}C-acetate PET[280,283] and one study of ^{18}F-fluorochlorine,[287] this pattern was not observed, due in part to higher sensitivity for more well-differentiated lesions.

KQ1.b. What is the comparative effectiveness of imaging-based surveillance strategies on intermediate outcomes such as diagnostic thinking?

No study compared effects of different imaging modalities or surveillance strategies on diagnostic thinking or clinical decisionmaking.

KQ1.c. What is the comparative effectiveness of imaging-based surveillance strategies on clinical and patient-centered outcomes?

One cluster randomized controlled trial (n=18816) conducted in China compared screening every 6 months with noncontrast US plus serum AFP versus no screening in persons 35 to 79 years of age (mean 42 years) with HBV infection (n=17250) or chronic hepatitis without HBV infection (n=1566) (Appendix H).[47] Technical details regarding the US methods used were not reported. Patients with an AFP >20 g/l or solid liver lesion on US underwent repeat testing; patients with repeatedly positive results underwent further diagnostic evaluation, including repeat US and CT or MRI "when necessary". Final diagnoses were based on liver biopsy or long-term followup. The trial was rated as high risk of bias; important methodological shortcomings included inadequate description of randomization or allocation concealment methods, unblended design, failure to report attrition, and failure to control for clustering affects (Appendix C). In addition, outcomes were based on physician reporting or data from the Shanghai Cancer Registry, but the completeness and accuracy of outcomes ascertainment could not be determined.

All screened patients underwent 5 to 10 cycles of screening; compliance with screening was 58 percent. The trial found screening associated with lower risk of HCC-related mortality (32 vs. 54 deaths, rate ratio 0.63, 95% CI 0.41 to 0.98) at 5-year followup. Screening was associated with a trend towards more HCC diagnoses (86 vs. 67, rate ratio 1.37, 95% CI 0.99 to 1.89), but also more Stage I (subclinical or early stage) cancers (52 vs. 0), with more patients undergoing surgical resection. All-cause mortality and harms were not reported.

One other randomized trial[296] compared screening versus no screening, but did not meet inclusion criteria because AFP testing was the primary mode of screening, with US only obtained to evaluate high AFP values. It found no difference between screening and no screening in risk of all-cause or HCC mortality.

Two trials compared different US screening intervals (Appendix H).[45,46] Technical details regarding the US methods used were not reported. One cluster randomized trial in Taiwan (n=744) found no difference between 4- versus 12-month intervals in risk of mortality after 4 years in patients with HBV or HCV infection (57% vs. 56%), even though more frequent

screening was associated with higher likelihood of early stage disease (37.5 vs. 6.7%, p=0.017).[46] The second trial (n=1278) in France and Belgium found no difference between 3- versus 6-month intervals in all-cause mortality in patients with cirrhosis related to alcohol use or viral hepatitis.[45]

KQ1.d. What are the adverse effects or harms associated with imaging-based surveillance strategies?

Two studies that met inclusion criteria reported harms associated with diagnostic imaging for HCC. One study reported 25 percent of patients (n=178) undergoing MRI experienced an adverse event following gadoxetic acid administration, with 56 events classified as mild and 6 as moderate.[131] There were two events classified as serious (anemia and hypotension); neither was considered related to the study drug. Twenty-one drug-related adverse events were reported in 10 percent of the patients, with nausea (1.7%) the most frequently reported event. One other study reported no clear differences between CT with contrast at 3 ml/s versus 5 ml/s in rate of overall adverse events (13% and 15%), discomfort (8% vs. 2%), or adverse events not related to contrast agents (5% vs. 3%).[189] No study reported rates of adverse events associated with use of microbubble contrast agents in US, and harms were not reported in randomized trials of screening with imaging.

Key Question 2. What is the comparative effectiveness of imaging techniques in diagnosing HCC among individuals in whom an abnormal lesion has been detected while undergoing surveillance for HCC or through the evolution of symptoms and abdominal imaging done for other indications?

Description of Included Studies

Forty-four studies[42,43,50,52,54,55,58,60,61,63,67,69,74,78-81,83,84,87,88,90,91,94,96,97,101-104,115,121,129,209-212,224,226,244,256,259,273,278] evaluated diagnostic accuracy of imaging tests in diagnosing HCC among individuals in whom an abnormal lesion has been detected and 15 studies[86,95,141,166,191,196,218,219,222,243,247,255,264,266,269] evaluated the accuracy of imaging tests for distinguishing HCC from another specific type of liver lesion.

No study compared effects of different imaging modalities or strategies on diagnostic thinking or on clinical or patient-centered outcomes. One study reported harms.[83]

Key Points

Test performance

- For evaluation of a previously identified lesion, using patients with HCC as the unit of analysis:
  - US with contrast: Sensitivity was 0.88 (95% CI 0.79 to 0.94, 8 studies) and specificity was 0.92 (95% CI 0.84 to 0.96, 5 studies), for a LR+ of 11 (95% CI 5.5 to 20) and LR- of 0.13 (95% CI 0.07 to 0.24).
  - US without contrast: Sensitivity was 0.78 (95% CI 0.72 to 0.83) in 2 studies; specificity was not reported in the studies.

- CT: Sensitivity was 0.85 (95% CI 0.76 to 0.91, 5 studies) and specificity was 0.92 (95% CI 0.86 to 0.96, 3 studies), for a LR+ of 11 (95% CI 5.7 to 22) and LR- of 0.17 (95% CI 0.10-0.27).
- MRI: Sensitivity was 0.76 (95% CI 0.62 to 0.86, 3 studies) and specificity was 0.87 (95% CI 0.70-0.95, 3 studies), for a LR+ of 5.9 (95% CI 2.5 to 14) and LR- of 0.28 (95% CI 0.18 to 0.43).

- For evaluation of a previously identified lesion, using HCC lesions as the unit of analysis:
  - US with contrast: Sensitivity was 0.86 (95% CI 0.79 to 0.91, 21 studies) and specificity was 0.93 (95% CI 0.87 to 0.96, 11 studies) for a LR+ of 12 (95% CI 6.3 to 21) and LR- of 0.15 (95% CI 0.10-0.23).
  - CT: Sensitivity was 0.80 (95% CI 0.67 to 0.88, 12 studies) and specificity was 0.89 (95% CI 0.29 to 0.99, 6 studies), for a LR+ of 6.9 (95% CI 0.53 to 91) and LR- of 0.23 (95% CI 0.13 to 0.40).
  - MRI: Sensitivity was 0.79 (95% CI 0.69 to 0.87, 13 studies) and specificity was 0.95 (95% CI 0.82 to 0.99, 12 studies), for a LR+ of 15 (95% CI 4.4 to 50) and LR- of 0.22 (95% CI 0.15 to 0.33).
  - PET: Sensitivity was 0.56 to 0.57 and specificity was 1.0 in two studies of FDG PET.
- For distinguishing HCC lesions from non-HCC hepatic lesions:
  - US with contrast: One study found US with sulfur hexafluoride contrast associated with a sensitivity of 0.94 (62/66) and a specificity of 0.68 (23/34) for distinguishing hypervascular HCC from focal nodular hyperplasia, using quantitative methods.
  - CT: Four studies evaluated accuracy of CT for distinguishing HCC from non-HCC lesions, but the non-HCC lesions varied in the studies, precluding strong conclusions.
  - MRI: Four studies reported inconsistent results for distinguishing small (<2 to 3 cm) hypervascular HCC lesions from hypervascular pseudolesions, with sensitivity of 0.47 and 0.52 in two studies, and 0.91 and 0.92 in the other two. Specificity was 0.93 or higher in all four studies. Five other studies evaluated accuracy of MRI for distinguishing HCC from other non-HCC lesions, but the non-HCC lesions varied in the studies, precluding strong conclusions.

- For direct (within-study) comparisons of imaging modalities, using patients with HCC as the unit of analysis:
  - US without contrast versus CT: Sensitivity was 0.78 (95% CI 0.70-0.85) versus 0.89 (95% CI 0.84 to 0.95), for a difference of -0.12 (95% CI -0.21 to -0.02), based on one study.
  - US with contrast versus CT: Sensitivity was 0.91 (0.85 to 0.95) versus 0.87 (95% CI 0.79 to 0.92), for a difference of 0.04 (95% CI -0.01 to 0.10), based on four studies.
  - MRI versus CT: Sensitivity was 0.81 (95% CI 0.70-0.92) versus 0.74 (95% CI 0.62 to 0.87), for a difference of 0.06 (-0.10 to 0.23), based on one study.

- Direct (within-study) comparisons of imaging modalities, using HCC lesions as the unit of analysis
  - US with contrast versus CT: Sensitivity was 0.94 (95% CI 0.89 to 0.97) versus 0.91 (95% CI 0.85 to 0.94), for a difference of 0.03 (95% CI -0.03 to 0.09), based on three studies.

- US with contrast versus MRI: Sensitivity was 0.79 (95% CI 0.65 to 0.94) versus 0.83 (95% CI 0.69 to 0.97), for a difference of -0.03 (95% CI -0.24 to 0.17), based on one study.
- MRI versus CT: One study found MRI associated with higher sensitivity (0.84, 95% CI 0.76 to 0.92 versus 0.62, 95% CI 0.52 to 0.72, for a difference of 0.22, 95% CI 0.09 to 0.35) but lower specificity (0.36, 95% CI 0.20-0.52 versus 0.72, 95% CI 0.58 to 0.87, for a difference of -0.36, 95% CI -0.58 to -0.15) than CT.

- Multiple imaging modalities
  - In four studies in which positive results with multiple modality imaging were defined as concordant typical findings for HCC on two imaging modalities, sensitivity was lower than with a single modality (difference in sensitivity ranged from 0.09 to 0.27), with no clear difference in specificity. In three studies in which positive results with multiple modality imaging were defined as typical findings for HCC on at least one of the imaging techniques, sensitivity was higher than with a single modality (increase in sensitivity ranged from 0.09 to 0.25), with no clear difference in specificity. One study found that a sequential imaging strategy, in which a second imaging test was only performed for indeterminant results on initial CT, increased sensitivity for HCC from 0.53 to 0.74 to 0.79.

- No study used explanted liver as the reference standard. There were no clear differences across imaging modalities in estimates of diagnostic accuracy in analyses stratified by use of different nonexplant reference standards.

- Sensitivity was substantially higher for lesions >2 cm in diameter than for lesions <2 cm in diameter, based on within-study comparisons:
  - US: Sensitivity was 0.91 (95% CI 0.53 to 0.99) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions < 2cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies.
  - CT and MRI: See Key Question 1.

- Sensitivity was higher for moderately- or poorly-differentiated HCC lesions than for well-differentiated HCC lesions, based on within-study comparisons:
  - US: Sensitivity was 0.84 (95% CI 0.64 to 0.94) for moderately- or poorly-differentiated HCC lesions and 0.43 (95% CI 0.21 to 0.69) for well-differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.22 to 0.59), based on four studies.
  - CT and MRI: See Key Question 1.

- Other factors
  - US: In two studies that directly compared US with versus without contrast, US with contrast was associated with sensitivity of 0.89 (95% CI 0.83 to 0.93) and US without contrast with a sensitivity of (0.39) 95% CI 0.32 to 0.47), for a difference in sensitivity of 0.50 (95% CI 0.41 to 0.58). Based on across-study comparisons, there were no clear differences in sensitivity between different US contrast agents; no study directly compared different contrast agents. There were no differences in sensitivity of US based on lesion depth (3 studies) or body mass index (2 studies).
  - CT: Evidence on effects of technical parameters (type of CT scanner, use of delayed phase imaging, section thickness) was limited by small numbers of studies with wide

confidence intervals and methodological limitations, precluding reliable conclusions. Two studies found no clear difference in sensitivity of CT for HCC in patients with versus without cirrhosis.

- MRI: There were no clear differences in estimates of sensitivity based on the type of MRI machine (3.0 T versus 1.5 T), type of contrast, use of delayed phase imaging, timing of delayed phase imaging, and section thickness. Estimates were similar when studies that used diffusion-weighted imaging were excluded.

Diagnostic Thinking

- No study compared effects of different imaging modalities or strategies on diagnostic thinking.

Clinical and Patient-centered Outcomes

- No study compared effects of different imaging modalities or strategies on clinical outcomes.

Harms

- One study of US (with and without contrast) and CT reported harms, but did not stratify results by imaging technique. The overall rate of adverse drug-related events was 10 percent, with all events classified as mild.

Detailed Synthesis

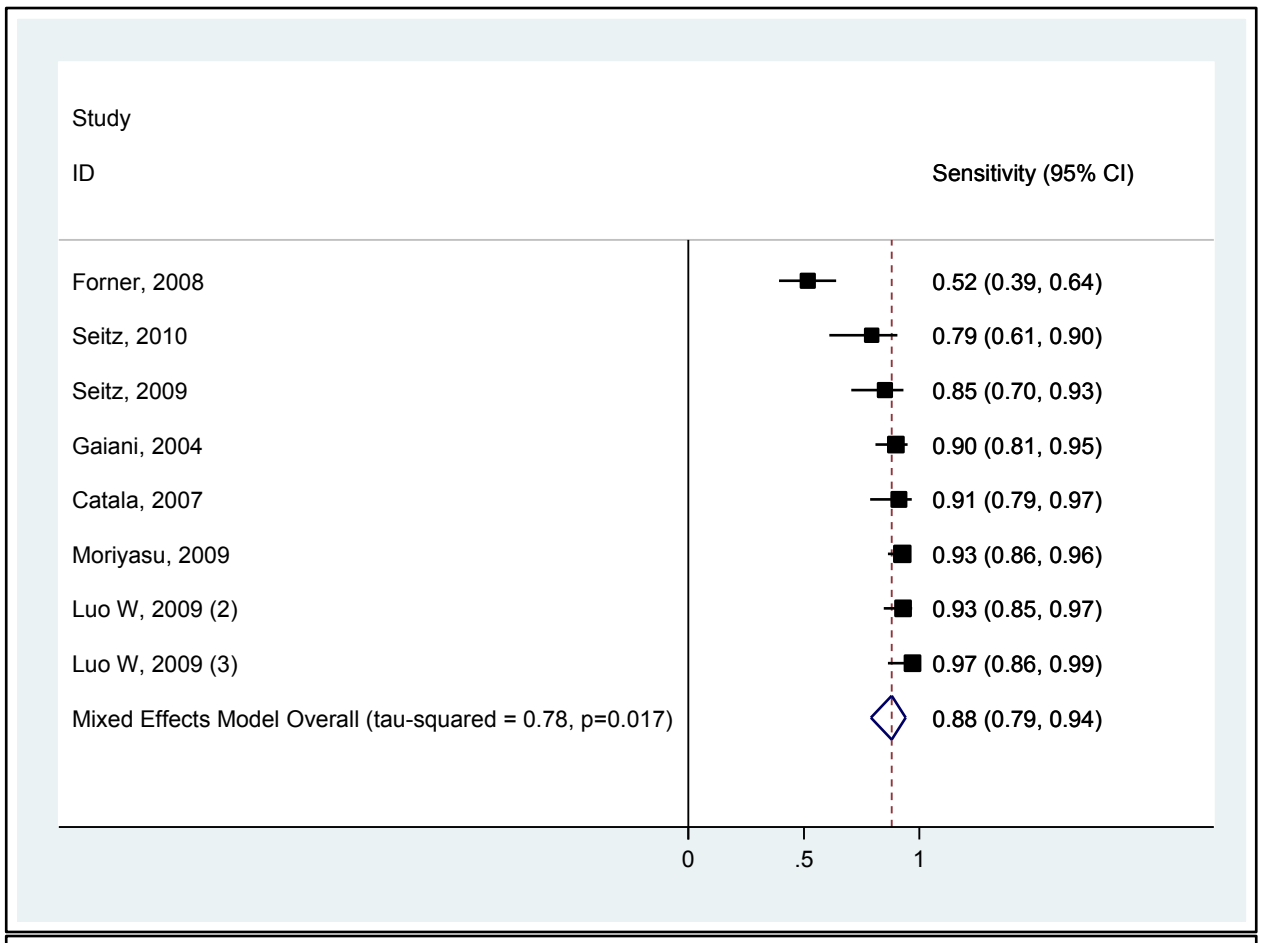
KQ2.a. What is the comparative test performance of imaging techniques for diagnosing HCC?

Ultrasound

For evaluation of a previously identified lesion, using patients with HCC as the unit of analysis, sensitivity of US with contrast was 0.88 (95% CI 0.79 to 0.94, 8 studies) and specificity was 0.92 (95% CI 0.84 to 0.96, 5 studies), for a LR+ of 11 (95% CI 5.5 to 20) and LR- of 0.13 (95% CI 0.07 to 0.24) (Table 6; Figure 23).[50,55,58,79,80,83,90,91] Using HCC lesions as the unit of analysis, sensitivity of US with contrast was 0.86 (95% CI 0.79 to 0.91, 21 studies) and specificity was 0.93 (95% CI 0.87 to 0.96, 11 studies) for a LR+ of 12 (95% CI 6.3 to 21) and LR- of 0.15 (95% CI 0.10-0.23) (Figures 24 and 25).[42,50,52,54,58,60,61,63,69,78,81,84,87,88,94,96,97,101-104] Sensitivity analyses based on study country, use of prospective design, use of Doppler, excluding high risk of bias studies, avoidance of case-control design, and interpretation of imaging blinded to the reference standard had little impact on estimates, and did not reduce heterogeneity.

One study found US with sulfur hexafluoride contrast associated with a sensitivity of 0.94 (62/66) and specificity of 0.68 (23/34) for distinguishing hypervascular HCC from focal nodular hyperplasia, based on quantitative analysis of US findings (Table 19).[86]

Figure 23. Test performance of ultrasound in evaluation of focal liver lesions for identification of patients with hepatocellular carcinoma


Study
ID
Sensitivity (95% CI)
Forner, 2008
0.52 (0.39, 0.64)
Seitz, 2010
0.79 (0.61, 0.90)
Seitz, 2009
0.85 (0.70, 0.93)
Gaiani, 2004
0.90 (0.81, 0.95)
Catala, 2007
0.91 (0.79, 0.97)
Moriyasu, 2009
0.93 (0.86, 0.96)
Luo W, 2009 (2)
0.93 (0.85, 0.97)
Luo W, 2009 (3)
0.97 (0.86, 0.99)
Mixed Effects Model Overall (tau-squared = 0.78, p=0.017)
0.88 (0.79, 0.94)
0
.5
1

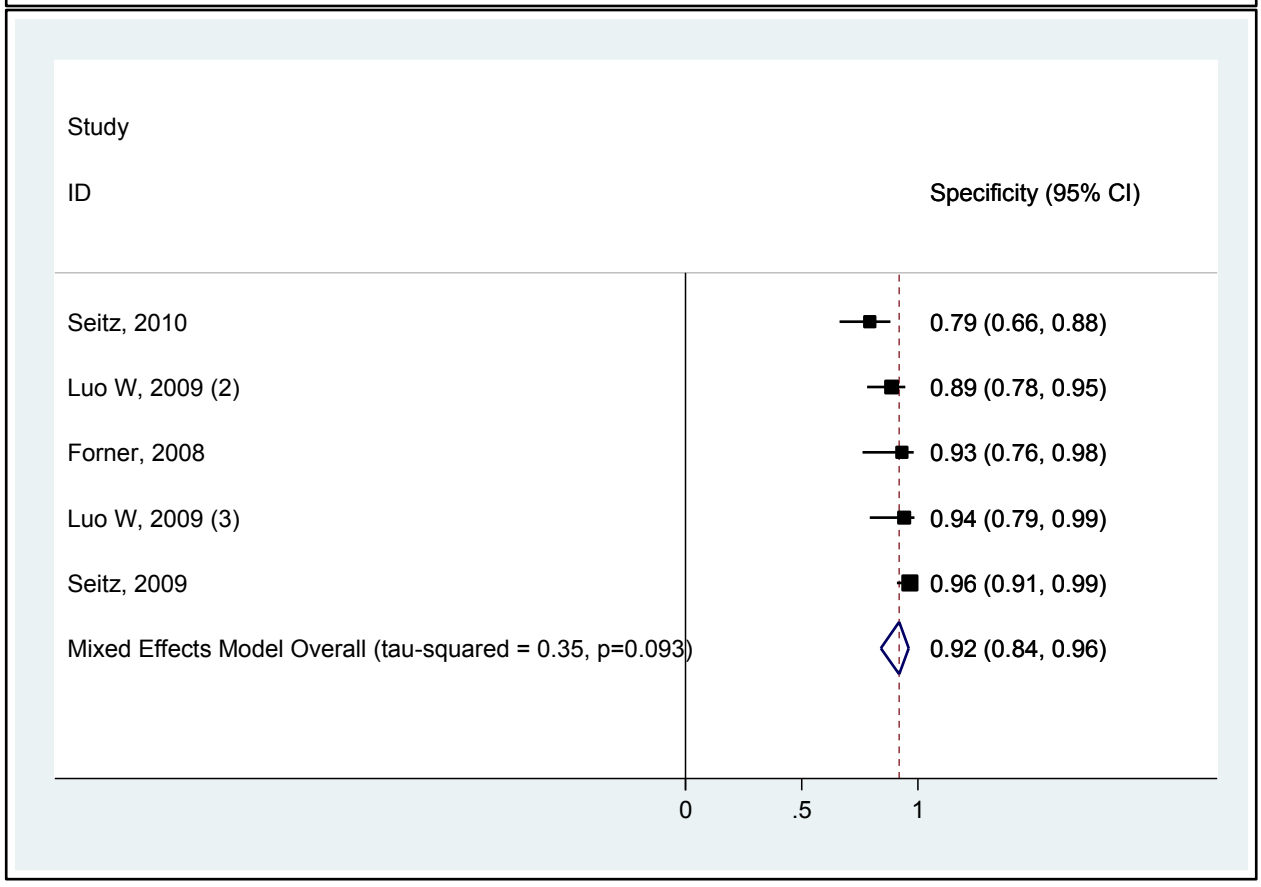

Study
ID
Specificity (95% CI)
Seitz, 2010
0.79 (0.66, 0.88)
Luo W, 2009 (2)
0.89 (0.78, 0.95)
Forner, 2008
0.93 (0.76, 0.98)
Luo W, 2009 (3)
0.94 (0.79, 0.99)
Seitz, 2009
0.96 (0.91, 0.99)
Mixed Effects Model Overall (tau-squared = 0.35, p=0.093)
0.92 (0.84, 0.96)
0
.5
1

Figure 24. Sensitivity of ultrasound with contrast for evaluation of focal liver lesions for identification of hepatocellular carcinoma lesions

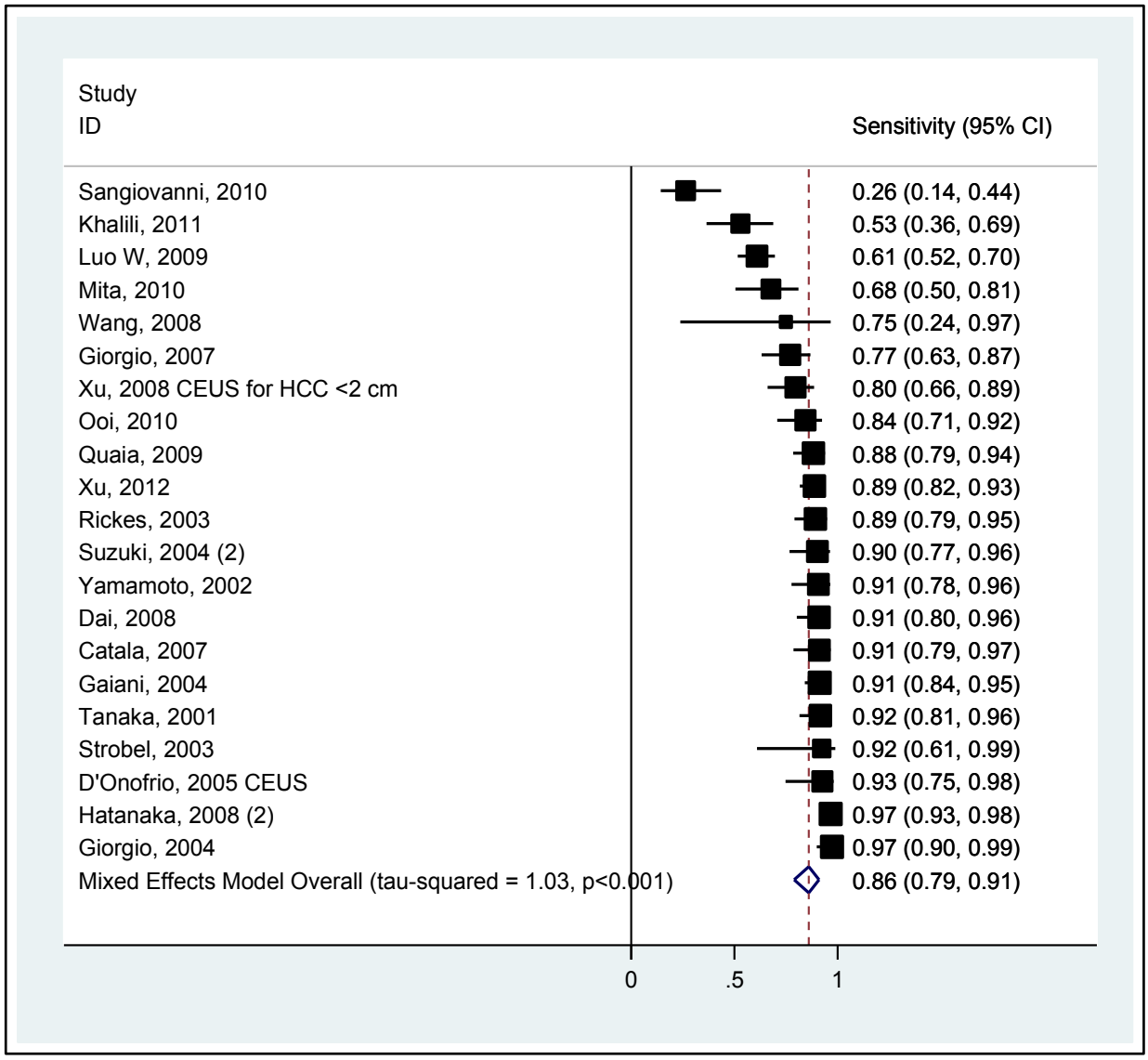
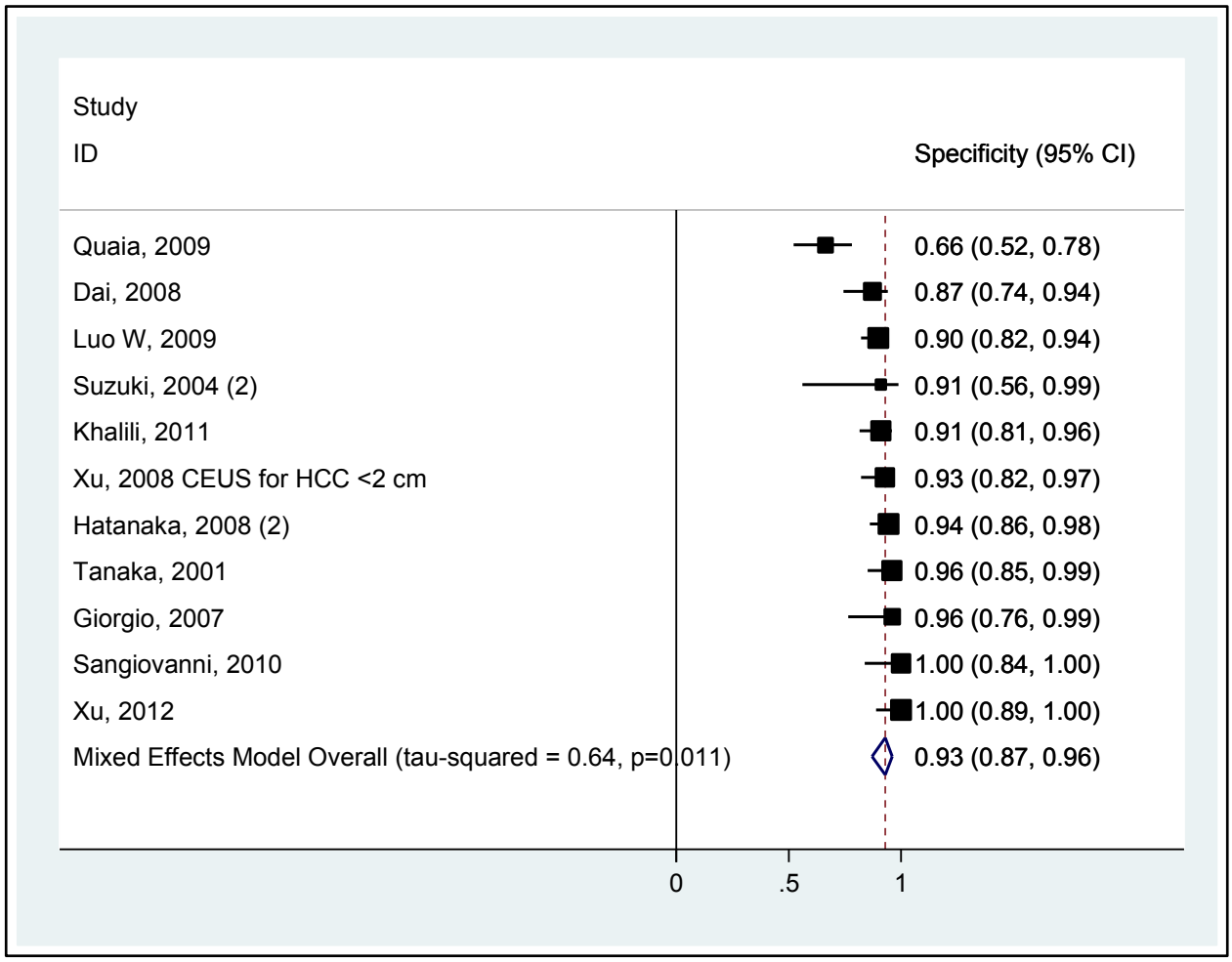


Figure 25. Specificity of ultrasound with contrast for evaluation of focal liver lesions for identification of hepatocellular carcinoma lesions

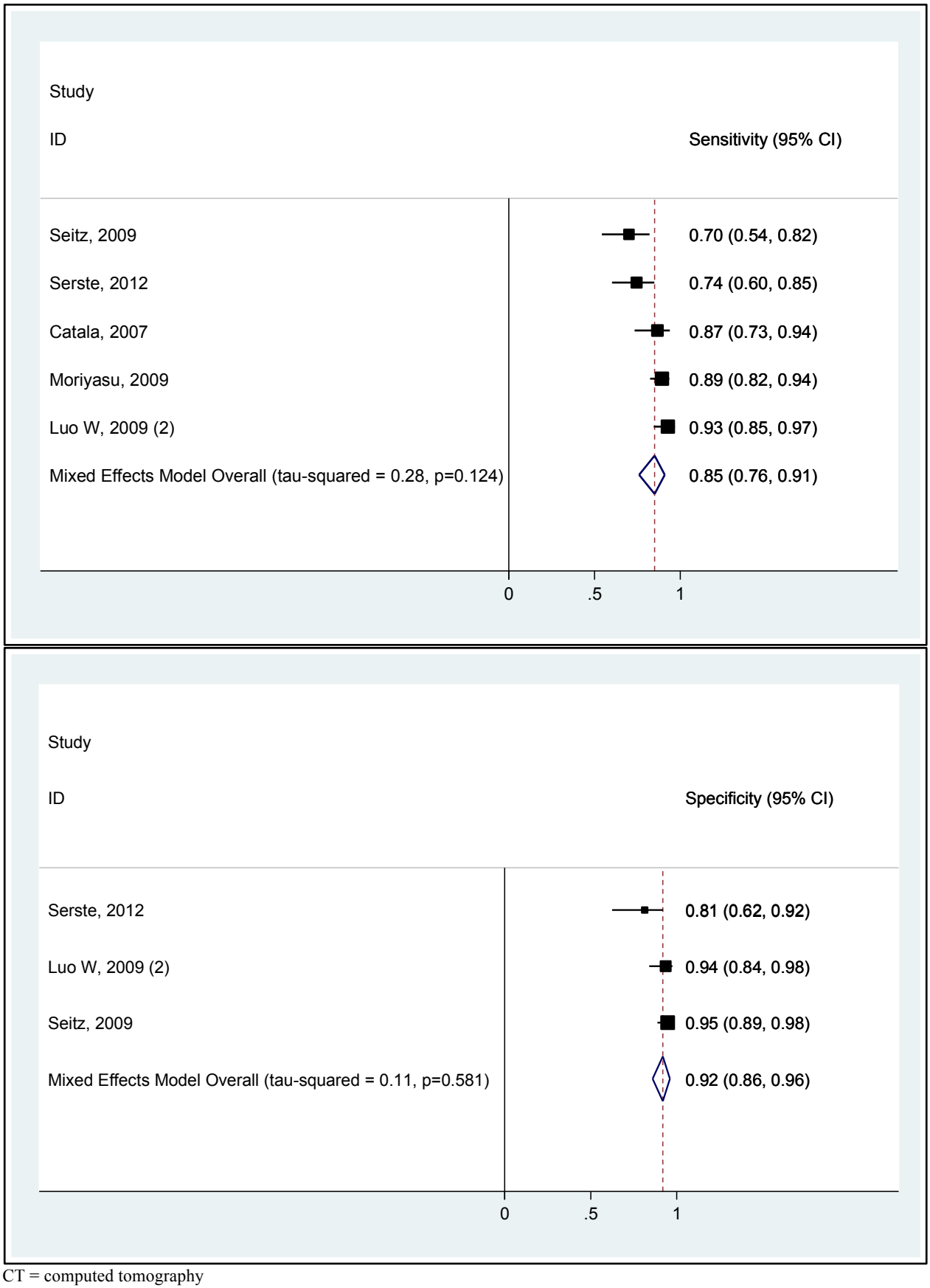


Computed Tomography

For evaluation of a previously identified lesion, using patients with HCC as the unit of analysis, sensitivity of CT was 0.85 (95% CI 0.76 to 0.91, 5 studies) and specificity was 0.92 (95% CI 0.86 to 0.96, 3 studies), for a LR+ of 11 (95% CI 5.7 to 22) and LR- of 0.17 (95% CI 0.10-0.27) (Table 7; Figure 26).[43,50,79,83,91] Using HCC lesions as the unit of analysis, sensitivity was 0.80 (95% CI 0.67 to 0.88, 11 studies) and specificity was 0.89 (95% CI 0.29 to 0.99, 5 studies), for a LR+ of 6.9 (95% CI 0.53 to 91) and LR- of 0.23 (95% CI 0.13 to 0.40) (Figures 27 and 28).[42,50,52,61,69,81,87,104,115,121,129] Excluding high risk of bias studies resulted in lower sensitivity (0.71, 95% 0.54 to 0.83, 8 studies) and specificity (0.66, 95% CI 0.11 to 0.97, 5 studies), for a LR+ of 2.1 (95% CI 0.39 to 11) and LR- of 0.45 (95% CI 0.23 to 0.87). Restricting the analysis to studies that were performed in the United States or Europe, used a prospective design, used a confidence rating scale, avoided a case-control design, or used blinded interpretation of imaging findings had little effect on estimates of sensitivity.

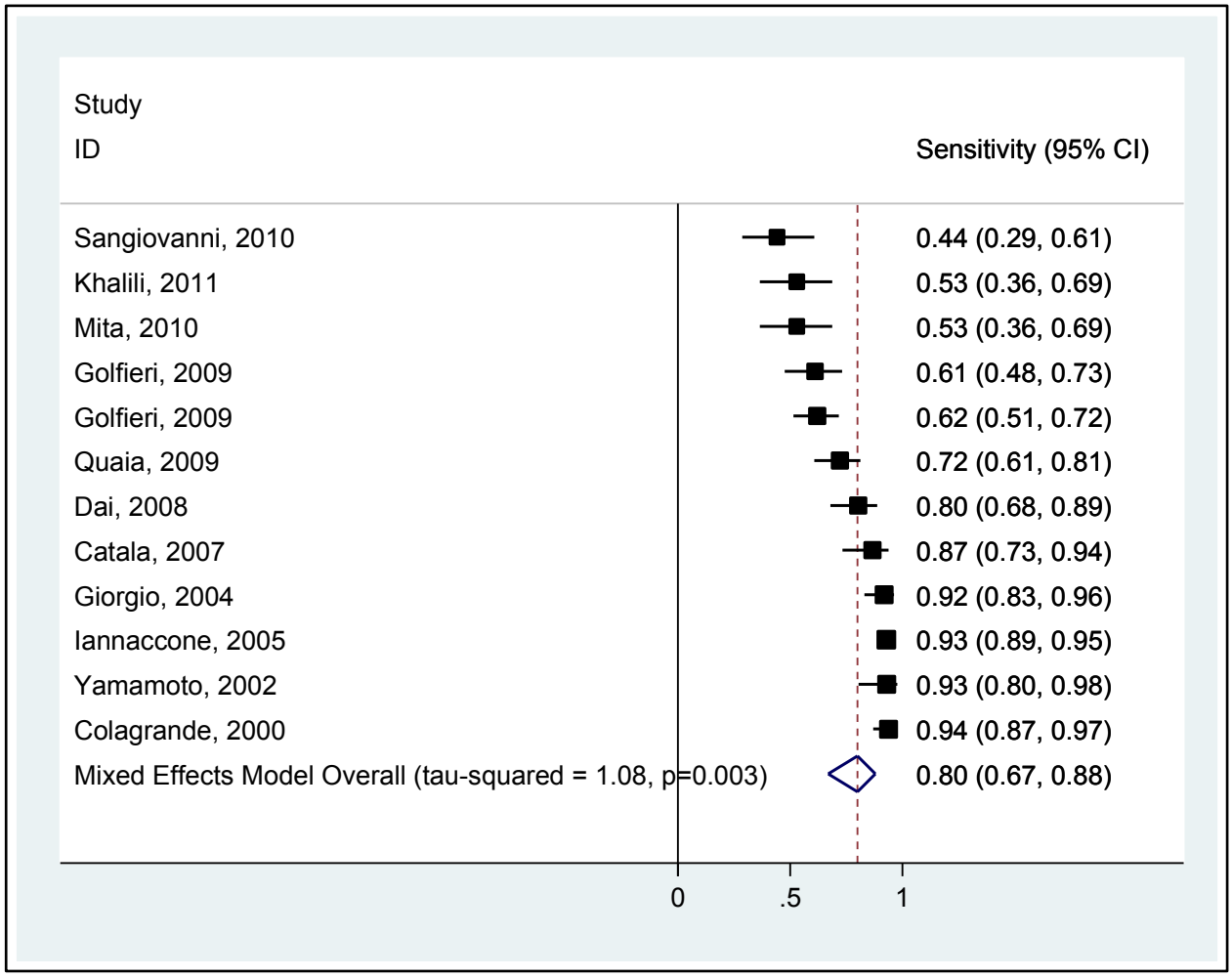
Four studies evaluated accuracy of CT for distinguishing HCC from non-HCC lesions (Table 19). The non-HCC lesions varied in the studies, precluding strong conclusions. In three studies, sensitivity ranged from 0.85 to 0.95 and specificity 0.87 to 1.0 for distinguishing HCC from hemangioma,[166] focal nodular hyperplasia,[196] or various non-HCC lesions.[141] In one study, CT was associated with a sensitivity of 0.54 (18/33) and specificity of 0.96 (26/27) for distinguishing hypervascular HCC lesions <2 cm from hypervascular pseudolesions.[191]

Figure 26. Test performance of CT in evaluation of focal liver lesions for identification of patients with hepatocellular carcinoma



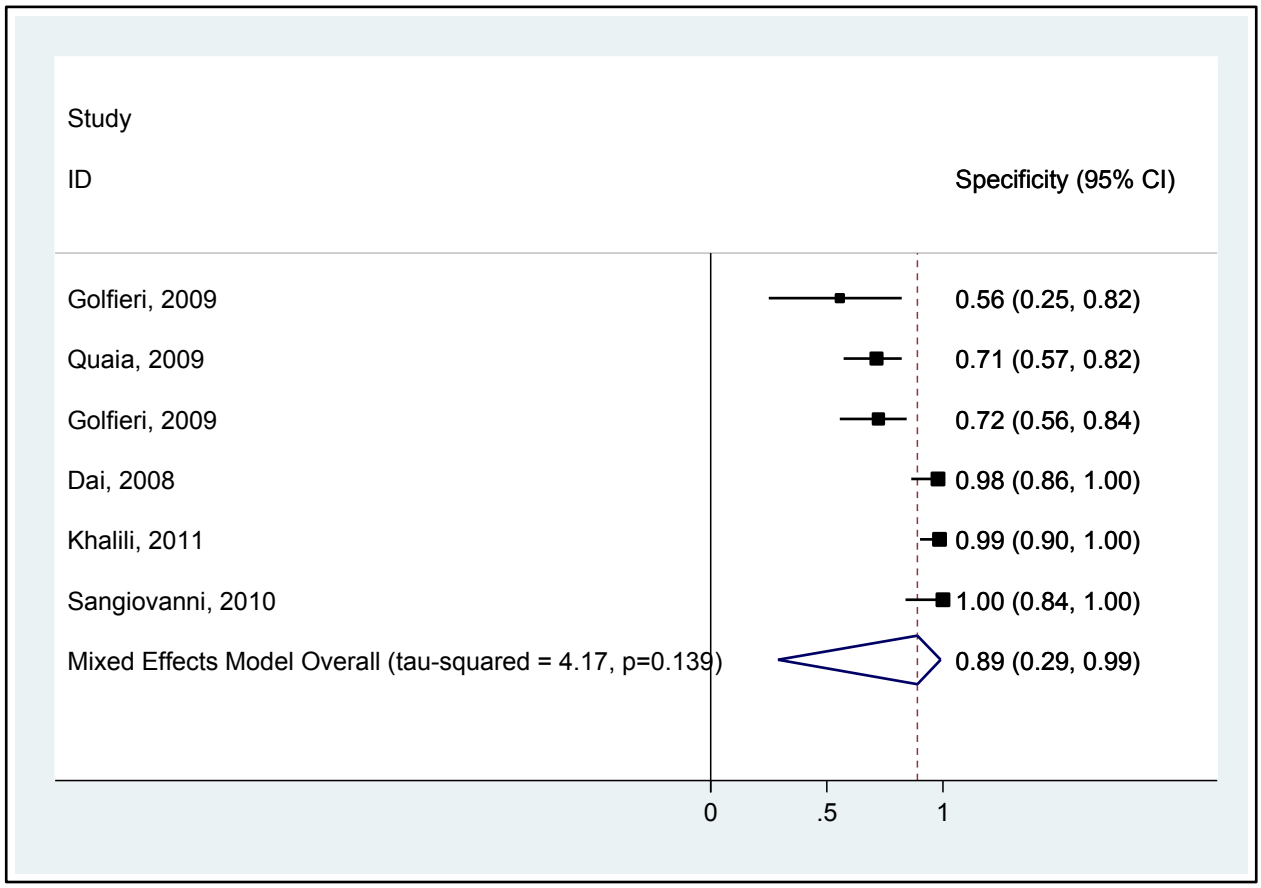
CT = computed tomography

Figure 27. Sensitivity of CT in evaluation of focal liver lesions for hepatocellular carcinoma lesions



CT = computed tomography

Figure 28. Specificity of CT in evaluation of focal liver lesions for hepatocellular carcinoma lesions



CT = computed tomography

Magnetic Resonance Imaging

For evaluation of a previously identified lesion, using patients with HCC as the unit of analysis, sensitivity of MRI was 0.76 (95% CI 0.62 to 0.86, 3 studies) and specificity was 0.87 (95% CI 0.70-0.95, 3 studies), for a LR+ of 5.9 (95% CI 2.5 to 14) and LR- of 0.28 (95% CI 0.18 to 0.43) (Table 8; Figures 29 and 30).[43,55,90]

Using HCC lesions as the unit of analysis, sensitivity was 0.79 (95% CI 0.69 to 0.87, 11 studies,[42,60,69,121,210,211,224,226,244,256,259] and specificity was 0.95 (95% CI 0.82 to 0.99, 10 studies),[42,60,69,121,210,224,226,244,256,259] for a LR+ of 15 (95% CI 4.4 to 50) and LR- of 0.22 (95% CI 0.15 to 0.33) (Figures 31 and 32). No study was rated high risk of bias. Excluding studies that were restricted to HCC lesions <2 cm increased the sensitivity to 0.92 (95% C I 0.87 to 0.95, 4 studies)[210,211,244,259] and excluding studies that were restricted to hypervascular HCC lesions decreased the sensitivity (0.69, 95% CI 0.45 to 0.80, 6 studies).[42,60,69,121,226,256] Restricting the analysis to studies that were performed in the United States or Europe, used a prospective design, used a confidence rating scale, avoided a case-control design, or used blinded interpretation of imaging findings had little effect on estimates of sensitivity.

Nine studies evaluated accuracy of MRI for distinguishing HCC from non-HCC lesions (Table 19). Four studies reported inconsistent results for distinguishing small (<2 to 3 cm) hypervascular HCC lesions from hypervascular pseudolesions, with sensitivity 0.47 and 0.52 in two studies,[218,222] and 0.91 and 0.92 in the other two.[191,247] Specificity was 0.93 or higher in all four studies. There was no clear pattern based on factors such as risk of bias, the diagnostic criteria applied, the reference standard, or the unit of analysis to account for the observed heterogeneity.

One study found MRI associated with poor specificity (0.15, 31/207) for distinguishing HCC lesions from cavernous hemangioma, with sensitivity of 0.88 (137/155), based on the absence of transient peritumoral enhancement.[269] Another study reported a sensitivity of 0.94 (31/33) and specificity of 0.82 (15/18) for distinguishing hypervascular HCC from hemangioma, based on quantitative evaluation of contrast-to-noise ratio.[219] Three other studies reported similar sensitivity (0.81 to 0.85) and specificity (0.42 to 0.65) for distinguishing HCC from dysplastic nodules[266] or various benign lesions.[243,264]

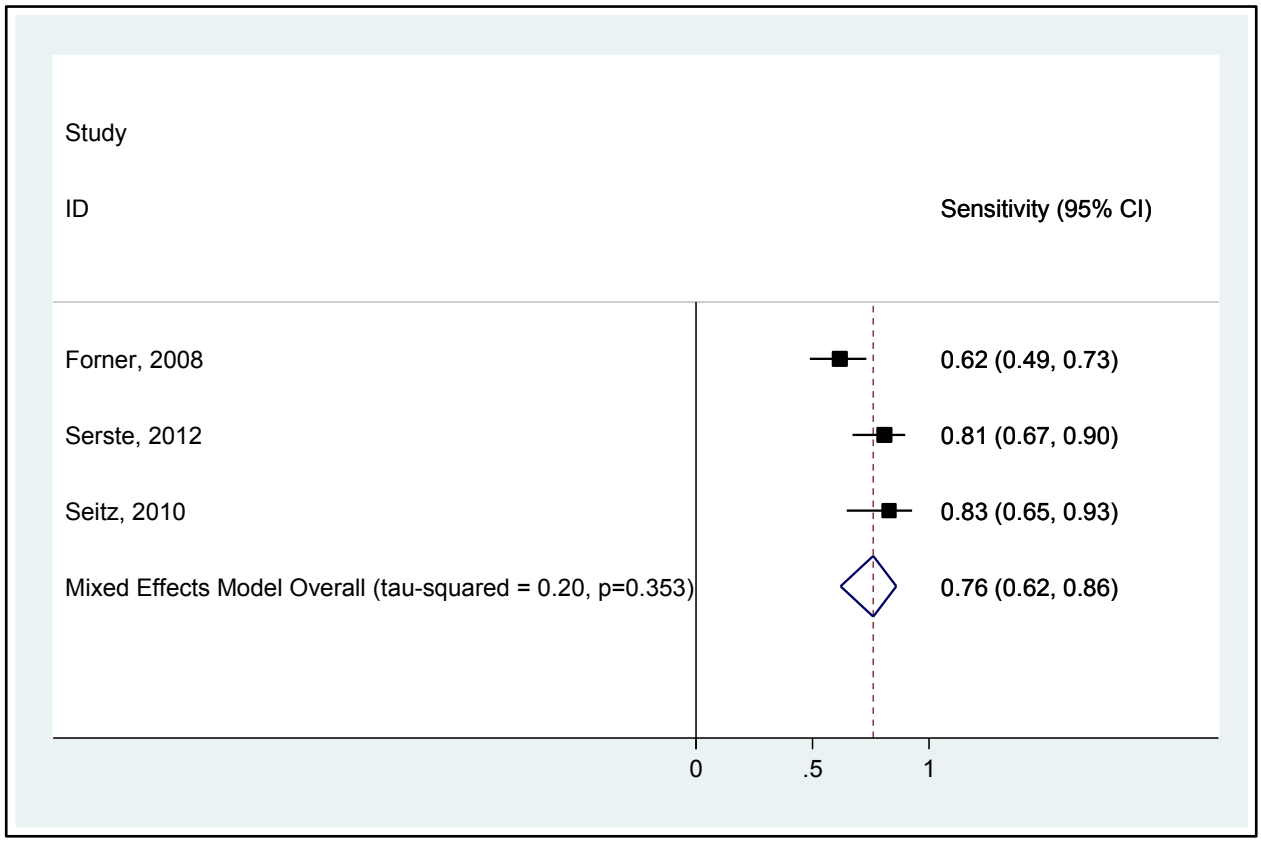
Positron Emission Tomography

For evaluation of a previously identified lesion, using patients with HCC as the unit of analysis, two studies reported similar sensitivity of FDG PET (0.56 to 0.57) and specificity of 1.0.[273,278]

Ultrasound versus Computed Tomography

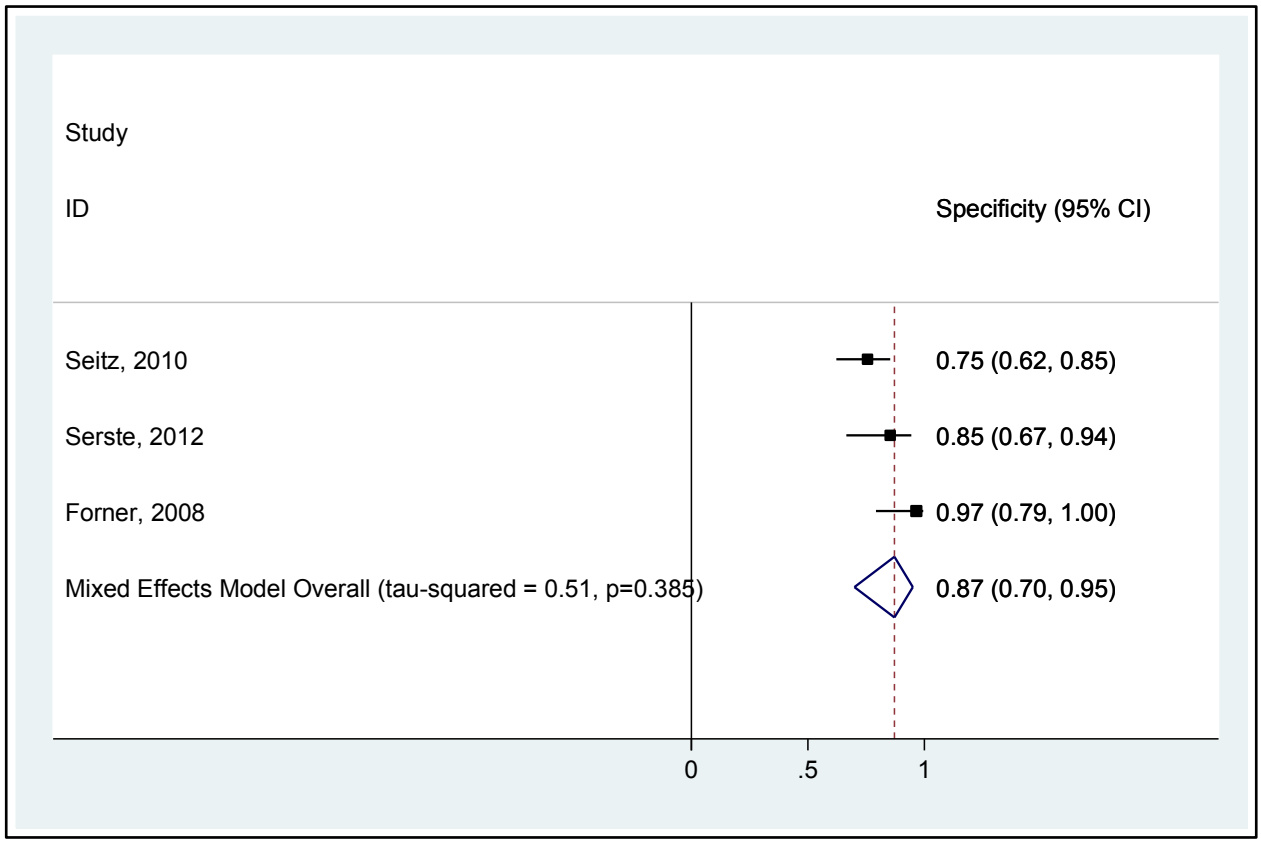
Using patients with HCC as the unit of analysis, based on four studies that directly compared diagnostic accuracy of imaging modalities, sensitivity was similar for US with contrast and CT (0.91, 95% CI 0.85 to 0.95 versus 0.87, 95% CI 0.79 to 0.92) (Table 20).[50,79,83,91] Using HCC lesions as the unit of analysis, US with contrast and CT were associated with similar sensitivity (0.94, 95% CI 0.89 to 0.97 versus 0.91, 95% CI 0.85 to 0.94, for a difference of 0.03, 95% CI -0.03 to 0.09), based on three studies.[50,61,104] There were also no clear differences between US with contrast and CT for HCC lesions <2 cm (0.78, 95% CI 0.61 to 0.89 versus 0.71, 95% CI 0.52 to 0.85, for a difference of 0.07, 95% CI -0.01 to 0.15), based on seven studies.[42,52,61,69,81,87,104] There was also no difference between US with contrast versus CT for well-differentiated lesions, based on two studies.[50,81]

Figure 29. Sensitivity of MRI in evaluation of focal liver lesions for identification of patients with hepatocellular carcinoma



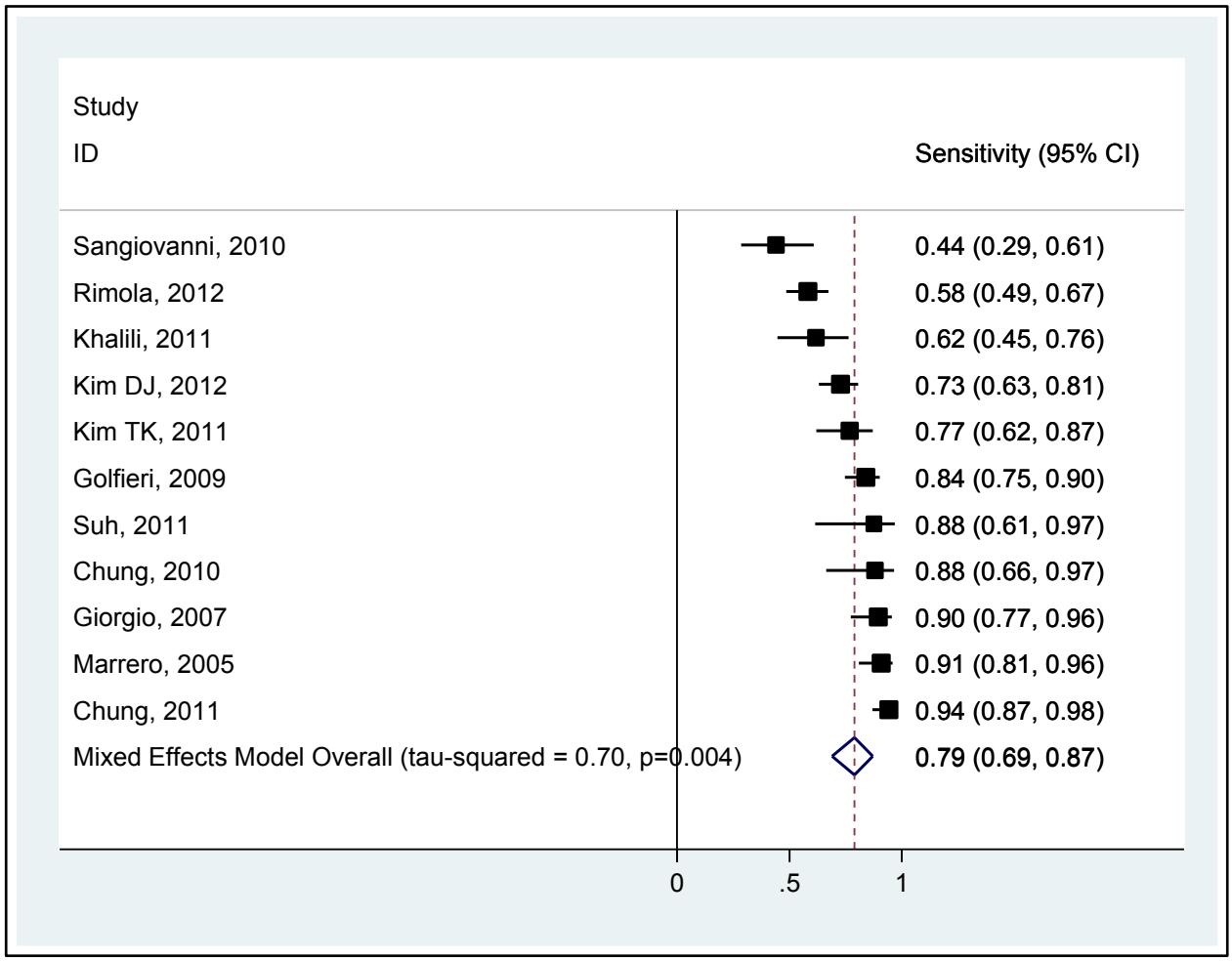
MRI = magnetic resonance imaging

Figure 30. Specificity of MRI in evaluation of focal liver lesions for patients with hepatocellular carcinoma



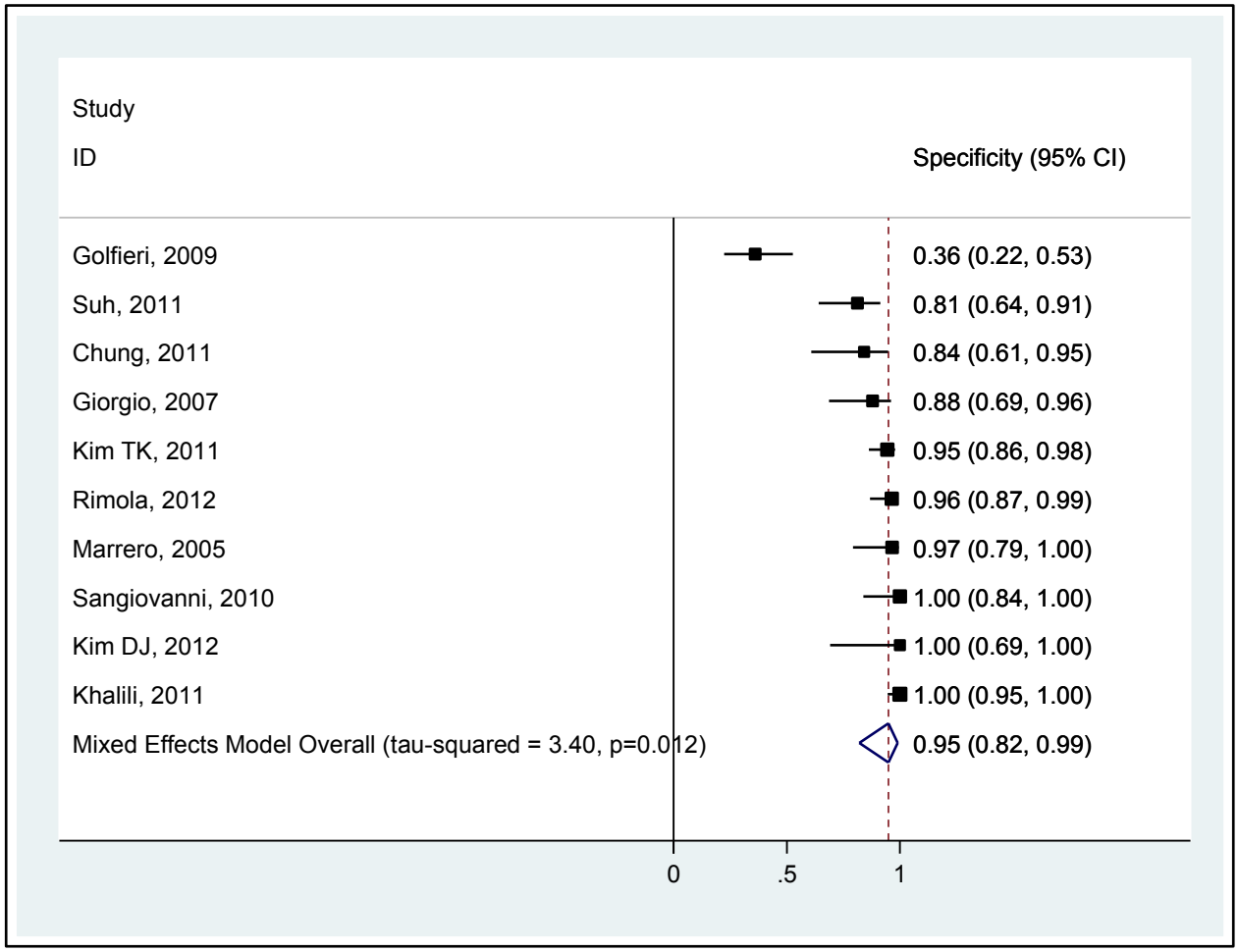
MRI = magnetic resonance imaging

Figure 31. Sensitivity of MRI in evaluation of focal liver lesions for hepatocellular carcinoma lesions



MRI = magnetic resonance imaging

Figure 32. Specificity of MRI in evaluation of focal liver lesions for hepatocellular carcinoma lesions



MRI = magnetic resonance imaging

Ultrasound versus Magnetic Resonance Imaging

Using patients with HCC as the unit of analysis, one study found no difference in sensitivity between US with contrast and MRI (0.79, 95% CI 0.65 to 0.94 versus 0.83, 95% CI 0.69 to 0.97).[90]

Magnetic Resonance Imaging versus Computed Tomography

Using patients with HCC as the unit of analysis, one study found no clear difference between MRI and CT in sensitivity or specificity.[43] Using HCC lesions as the unit of analysis, one study found MRI associated with higher sensitivity (0.84, 95% CI 0.76 to 0.92 vs. 0.62, 95% CI 0.52 to 0.72, for a difference of 0.22, 95% CI 0.09 to 0.35) but lower specificity (0.36, 95% CI 0.20-0.52 vs. 0.72, 95% CI 0.58 to 0.87, for a difference of -0.36, 95% CI -0.58 to -0.15) than CT.[121]

Multiple Imaging Modalities

Seven studies compared diagnostic performance of single versus multiple modality imaging for diagnosis of HCC (Table 11).[42,43,52,55,69,87,121] Five reported diagnostic accuracy for small (<2 or <3 cm) HCC lesions.[42,55,69,87,121] In four studies in which positive results with multiple modality imaging were defined as concordant typical findings for HCC on two imaging modalities, sensitivity was lower than with a single modality (decrease insensitivity ranged from 0.09 to 0.27), with no clear difference in specificity.[42,43,55,69] In three studies in which positive results with multiple modality imaging were defined as typical findings for HCC on at least one of the imaging techniques, sensitivity was higher than with a single modality (increases in sensitivity ranged from 0.09 to 0.25), with no clear difference in specificity.[42,43,87] One study found that a sequential imaging strategy in which a second imaging test was only performed for indeterminant results on initial CT increased sensitivity from 0.53 to 0.74 to 0.79.[69] Two other studies also found multiple imaging modalities associated with higher sensitivity than a single technique, but did not clearly describe criteria used to define a positive result with multiple modality imaging.[52,121] There were too few studies to evaluate the comparative diagnostic performance of different combinations of imaging modalities.

KQ2.a.i. How is a particular technique's test performance modified by use of various reference standards?

Ultrasound

No study evaluated diagnostic accuracy of US for evaluation of a previously identified lesion using explanted livers as the reference standard. There were no clear differences in sensitivity for nonexplant reference standards (histopathological, imaging/clinical criteria, or mixed), based on pooled sensitivity (range 0.77 to 0.91) with wide confidence intervals, using either patients with HCC or HCC lesions as the unit of analysis (Table 6).

Computed Tomography

No studies of CT for evaluation of previously identified lesions used explanted livers as the reference standard. Accuracy was similar for studies that used a nonexplant histopathological reference standard or a mixed (histological with clinical/imaging criteria) reference standard (Table 7).

Magnetic Resonance Imaging

No studies of MRI for evaluation of previously identified lesions used explanted livers as the reference standard. Sensitivity was somewhat lower in studies that used a nonexplant histopathological reference standard (0.72, 95% CI 0.53 to 0.85, 4 studies)[42,121,244,256] than a mixed (histological with clinical/imaging criteria) reference standard (0.83, 95% CI 0.72 to 0.91, 7 studies),[60,69,210,211,224,226,259] but confidence intervals were wide and overlapped (Table 8).

KQ2.a.ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?

Ultrasound

Studies of US reported higher sensitivity with contrast than without contrast when using patients with HCC as the unit of analysis (0.87, 95% CI 0.78 to 0.93, 9 studies vs. 0.78, 95% CI 0.72 to 0.83, 2 studies) or HCC lesions as the unit of analysis (0.86, 95% CI 0.77 to 0.92, 12 studies vs. 0.62, 95% CI 0.18 to 0.93, 4 studies) (Table 6). Two studies that directly compared US with versus without contrast using HCC lesions as the unit of analysis found US with contrast associated with higher sensitivity (0.89, 95% CI 0.83 to 0.93) than US without contrast (0.39, 95% CI 0.32 to 0.47), for a difference in sensitivity of 0.50 (95% CI 0.41 to 0.58) (Table 12). Based on patients as the unit of analysis, sensitivity was somewhat higher in studies that used perflubutane contrast (0.94, 95% CI 0.87 to 0.98, 3 studies)[78,79,83] than for studies that used sulfur hexafluoride contrast (0.82, 95% CI 0.69 to 0.90, 5 studies),[50,55,58,90,91] but using HCC lesions as the unit of analysis, there was no difference between sulfur hexafluoride (0.86, 95% CI 0.77 to 0.92, 12 studies),[42,50,52,54,58,60,61,84,87,101-103] perflubutane (0.82, 95% CI 0.51 to 0.95, 3 studies),[63,78,81] and galactose (0.90, 95% CI 0.86 to 0.94, 4 studies).[88,96,97,104] No study directly compared different types of contrast agents.

One study using patients with HCC as the unit of analysis compared US with contrast versus without contrast (sensitivity 0.93, 95% CI 0.88 to 0.97 vs. 0.78, 95% CI 0.70-0.85, for a difference of 0.15, 95% CI 0.06 to 0.23), but results were potentially confounded by use of Doppler in the contrast group.[83] Two studies using HCC lesions as the unit of analysis that directly compared US with versus without Doppler found no difference in sensitivity.

In 14 studies that directly compared accuracy of US stratified by HCC lesion size, sensitivity was greater for lesions >2 cm (0.91, 95% CI 0.53 to 0.99) than for lesions <2 cm (0.49, 95% CI 0.31 to 0.67), for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51). In four studies that directly compared accuracy of US stratified by degree of tumor differentiation, sensitivity was greater for moderately- or poorly-differentiated HCC lesions (0.84, 95% CI 0.64 to 0.94) than for well-differentiated lesions (0.43, 95% CI 0.21 to 0.69), for an absolute difference in sensitivity of 0.40 (95% CI 0.22 to 0.59).[50,57,64,81] There were no differences in sensitivity of US based on lesion depth (3 studies)[62,103,104] or body mass index (2 studies).[62,71]

Computed Tomography

Using HCC lesions as the unit of analysis, studies of nonmultidetector CT reported higher sensitivity (0.86, 95% CI 0.77 to 0.92, 6 studies) than studies of multidetector CT (sensitivity 0.57, 95% CI 0.37 to 0.74, 3 studies [≥8 rows] and 0.62, 95% CI 0.31 to 0.85, 2 studies [<8 rows]), studies without delayed phase imaging reported higher sensitivity (0.94, 95% CI 0.79 to 0.98, 2 studies) than studies with delayed phase imaging (sensitivity 0.75, 95% CI 0.62 to 0.84, 10 studies), and studies with a section thickness >5 mm reported higher sensitivity (0.91, 95 5CI 0.77 to 0.97, 2 studies) than studies with section thickness ≤5 mm (0.69, 95% CI 0.55 to 0.80, 8

studies) (Table 7). However, confidence intervals were wide and most studies that used methods not meeting minimum technical standards (nonmultidetector CT, no delayed phase imaging, or section thickness >5 mm) were rated high risk of bias. In two studies that directly compared sensitivity using a section thickness of 7.5 mm versus 5.0 mm, there was no clear difference (sensitivity 0.64, 95% CI 0.58 to 0.70 vs. 0.72, 95% CI 0.64 to 0.78, for a difference of -0.07, 95% CI -0.17 to 0.02) (Table 14).

Two studies found no clear difference in sensitivity of CT for HCC in patients with versus without cirrhosis. Effects of lesion size and tumor differentiation on accuracy are presented in the results for Key Question 1.

Magnetic Resonance Imaging

Using HCC lesions as the unit of analysis, sensitivity was higher for studies that evaluated 3.0 T MRI (0.89, 95% CI 0.64 to 0.97, 2 studies)[210,259] than studies that evaluated 1.5 T MRI (0.77, 95% CI 0.66 to 0.86, 11 studies) (Table 8).[42,60,69,121,211,224,226,244,256] However, confidence intervals were wide and overlapped. There was no clear difference in estimates of sensitivity based on type of contrast, use of delayed phase imaging, and timing of delayed phase imaging. Only one study evaluated MRI with enhanced section thickness of >5 mm.[210,259] Estimates were similar when studies that used diffusion-weighted imaging were excluded. Effects of lesion size and tumor differentiation on accuracy are presented in the results for Key Question 1.

KQ2.b. What is the comparative effectiveness of the various imaging techniques on intermediate outcomes such as diagnostic thinking and use of additional diagnostic procedures such as fine-needle or core biopsy?

No study evaluated the comparative effectiveness of different imaging techniques on outcomes such as diagnostic thinking and use of additional diagnostic procedures.

KQ2.c. What is the comparative effectiveness of the various imaging techniques on clinical and patient-centered outcomes?

No study evaluated the comparative effectiveness of different imaging techniques on clinical and patient-centered outcomes.

KQ2.d. What are the adverse effects or harms associated with imaging-based diagnostic strategies?

One study of US (with and without contrast) and CT reported harms, but did not stratify results by imaging technique. The overall rate of adverse drug-related events was 10 percent, with all events classified as mild.[83]

Key Question 3. What is the comparative effectiveness of imaging techniques in staging HCC among patients diagnosed with HCC?

Description of Included Studies

Six studies reported test performance of various imaging techniques for staging of patients with HCC based on TNM criteria.[41,56,109,113,114,198] Ten studies reported test performance of PET for detection of metastatic disease.[137,156,174,275,283-285,290,291,294]

Seven studies reported effects of imaging on transplant decisions[75,92,109,114,164,187,192] and one study reported comparative effects of imaging on clinical and patient-centered outcomes.[137,156,174,275,283-285,290,291,294] No study reported harms associated with imaging for HCC staging.

Key Points

Test performance

- For staging, using TNM criteria, using explanted liver or surgical resection reference standard:
  - CT: The proportion correctly staged ranged from 28 percent to 58percent, the proportion overstaged from 2 percent to 27percent, and the proportion understaged from 25 percent to 52 percent, based on six studies.
  - MRI: The proportion correctly staged were 10 percent and 31 percent, the proportion overstaged 10 percent and 31 percent, and the proportion understaged 29 percent and 31 percent, based on two studies.
  - PET: One study found 26 percent of patients were correctly staged with FDG PET and 91 percent with ^{11}C-choline PET.
  - MRI versus CT: Two studies reported similar staging accuracy.
- For identification of metastatic disease, using patients with metastatic HCC as the unit of analysis:
  - PET: Sensitivity of FDG PET was 0.85 (95% CI 0.71 to 0.93, 6 studies) and specificity was 0.93 (95% CI 0.89 to 0.95, 5 studies), for a LR+ of 11 (95% CI 7.8 to 17) and LR- of 0.16 (95% CI 0.08 to 0.33). One study that directly compared sensitivity of FDG PET to ^{11}C-acetate PET reported comparable sensitivity (0.79 vs. 0.71), though sensitivity was higher when both tracers were used (0.98).
  - PET/CT versus CT: Three studies found no difference in sensitivity (0.82, 95% CI 0.61 to 0.93 vs. 0.85, 95% CI 0.66 to 0.95).
- For identification of metastatic disease, using metastatic HCC lesions as the unit of analysis:
  - PET: Sensitivity of FDG PET was 0.82 (95% CI 0.72 to 0.90, 5 studies). One study that directly compared sensitivity of FDG to ^{11}C-acetate PET reported comparable sensitivity (0.86 vs. 0.77, respectively).
- Evidence was insufficient to determine effects of different reference standards on test performance:
  - For accuracy of staging using TNM criteria, all but one study used explanted livers as the reference standard.

- For accuracy of PET for identifying metastatic HCC, five of the six studies that used patients with metastatic HCC as the reference standard used mixed histological and imaging/clinical criteria as the reference standard. For studies that used metastatic HCC lesions as the unit of analysis, different reference standards were each evaluated in only one or two studies.

- Effects of patients, tumor, or technical factors on test performance:
  - For accuracy of staging using TNM criteria, no study evaluated effects of patient-level characteristics or other factors on accuracy of imaging techniques for staging.
  - For identifying metastatic HCC, estimates for sensitivity were too imprecise to determine how use of PET versus PET/CT affected test performance. In one study that directly compared sensitivity of PET versus PET/CT for identifying metastatic HCC lesions, there was no clear difference in sensitivity.
  - Four studies of PET with FDG found sensitivity increased as lesion size increased, but the number of lesions <1 cm was small (total of 20).
  - Eight studies reported test performance of FDG PET stratified by location of metastasis. In most studies, sensitivity was higher for lymph and bone metastasis than for lung metastasis, but samples were small, precluding strong conclusions.

Diagnostic Thinking

- Transplant eligibility, using Milan criteria
  - CT: The proportion correctly assessed for transplant eligibility ranged from 40 percent to 96 percent. Three studies reported the proportion of patients who met transplant criteria based on CT but exceeded criteria based on the reference standard was 3.5 to 7.8 percent, based on three studies. Two studies found that 2.3 percent and 16 percent of patients who underwent transplantation based on Milan criteria had no HCC lesions on examination of explanted livers.
  - CT versus MRI: One study reported similar accuracy.
  - PET versus CT: One study found ^{11}C-choline PET more accurate than CT (95% vs. 40%).
  - MRI versus CT: One study reported that the proportion of decisions to perform resection or ablative therapies that were classified as correct were similar for MRI (90% and 90%, respectively) and CT (80% and 77%, respectively).

Clinical and Patient-centered Outcomes

- US with contrast versus US without contrast plus CT: One cohort study found that contrast enhanced US identified more small (≤2 cm) HCC lesions than noncontrast US plus CT (36 vs. 31), and was associated with a higher complete necrosis rate following ablation (92% or 106/115 vs. 83% or 93/112 lesions, p=0.036), but was rated high risk of bias.

Harms

- No evidence.

Detailed Synthesis

KQ3.a. What is the comparative test performance of imaging techniques to predict HCC tumor stage?

Accuracy of Imaging for Staging

Six studies evaluated accuracy of imaging techniques for staging using TNM criteria (six studies) or BCLC (one study) criteria (Table 21).[41,56,109,113,114,198] Five studies used explanted liver as the reference standard and the sixth[114] used an explanted liver or surgical resection. CT was evaluated in six studies, MRI in two studies, and PET in one study. For CT, the proportion correctly staged ranged from 28 percent to 58 percent, the proportion overstaged from 2 percent to 27 percent, and the proportion understaged from 25 percent to 52 percent.[41,56,109,113,114,198] For MRI, the proportion correctly staged were 10 percent and 31 percent, the proportion overstaged 10 percent and 31 percent, and the proportion understaged 29 percent and 31 percent.[56,113] One study found 26 percent of patients correctly staged with FDG PET and 91 percent with ^{11}C-choline PET.[114]

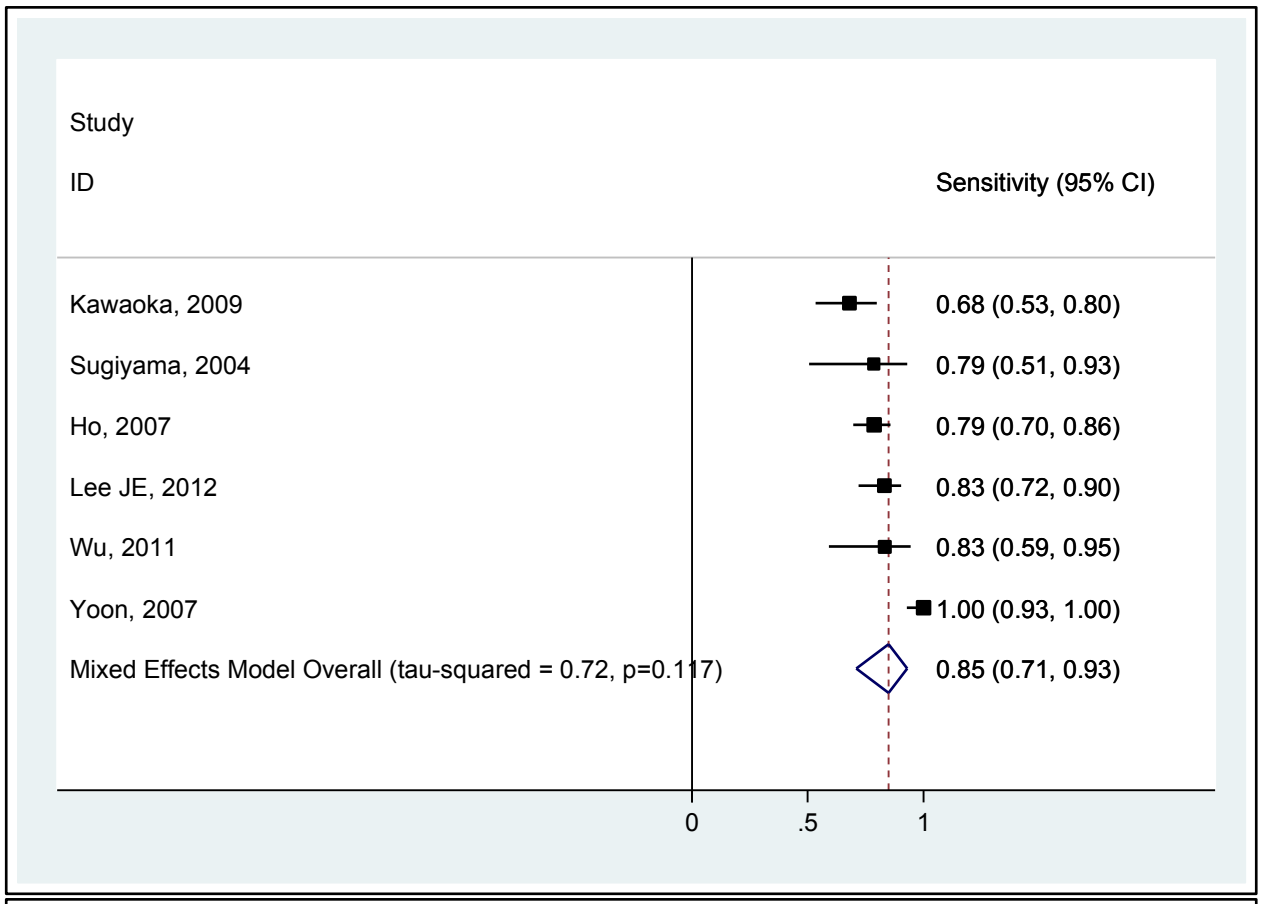
Two studies that directly compared staging accuracy of imaging modalities found similar staging accuracy for MRI versus CT.[56,113]

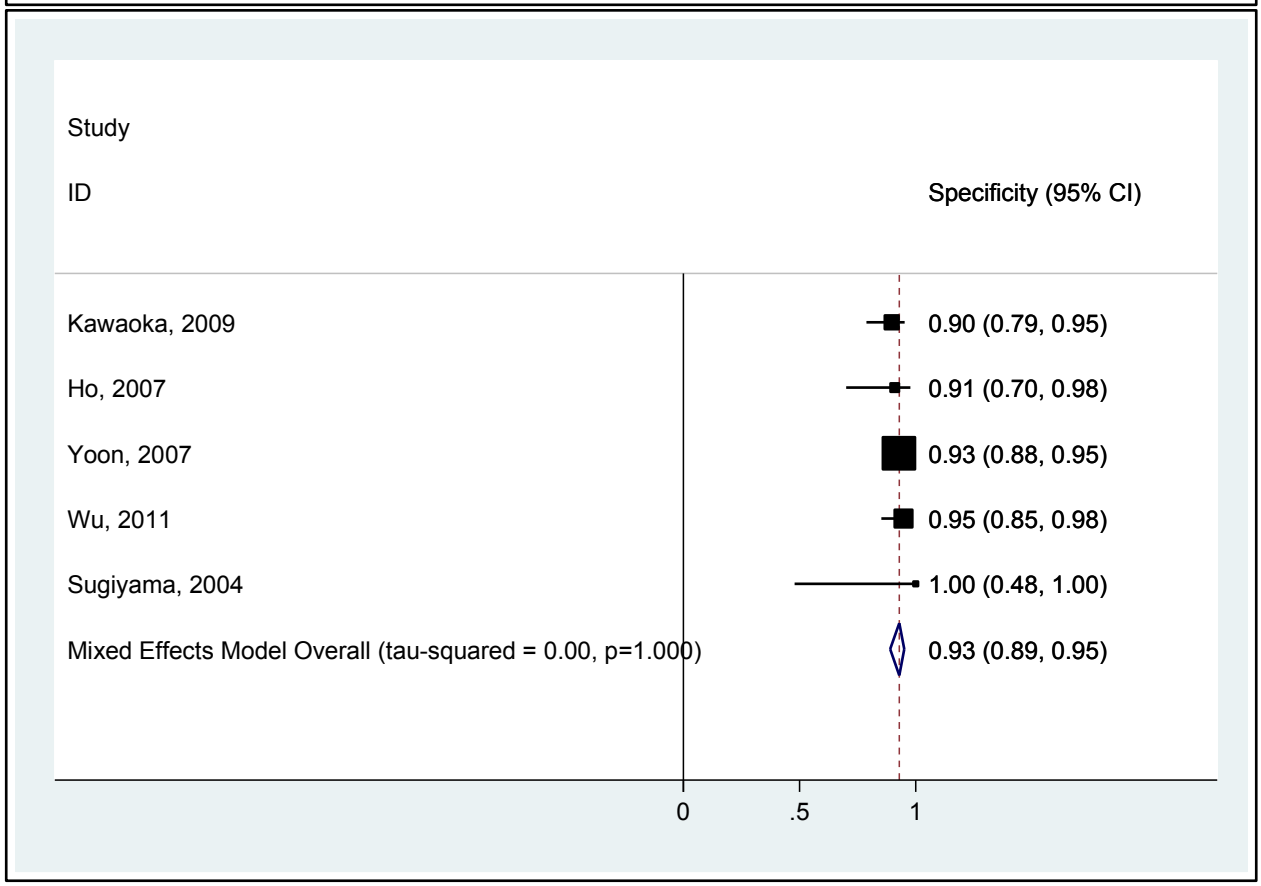
PET for Detection of Metastatic Hepatocellular Carcinoma Disease

Using patients with metastatic HCC as the unit of analysis, sensitivity of FDG PET was 0.85 (95% CI 0.71 to 0.93, 6 studies) and specificity was 0.93 (95% CI 0.89 to 0.95, 5 studies), for a LR+ of 11 (95% CI 7.8 to 17) and LR- of 0.16 (95% CI 0.08 to 0.33) (Table 9; Figure 33).[137,156,275,285,291,294] Estimates were similar when high risk of bias studies were excluded and when the analysis was restricted to studies that used a prospective design. All studies were conducted in Asia. One study that directly compared sensitivity of FDG PET to ^{11}C-acetate PET reported comparable sensitivity (0.79 vs. 0.71), though sensitivity was higher when both tracers were used (0.98).[275]

Using metastatic HCC lesions as the unit of analysis, sensitivity of FDG PET was 0.82 (95% CI 0.72 to 0.90, 5 studies) (Figure 34).[174,279,283,285,290] No study reported specificity. All studies except one were rated high risk of bias. In the one moderate risk of bias study, sensitivity was 0.86 (95% CI 0.70-0.95).[283] All studies were conducted in Asia except for one small study (n=5) conducted in the United States.[290] One study that directly compared sensitivity of FDG to ^{11}C-acetate PET reported comparable sensitivity (0.86 vs. 0.77, respectively) (Table 16. PET direct comparisons).[283]

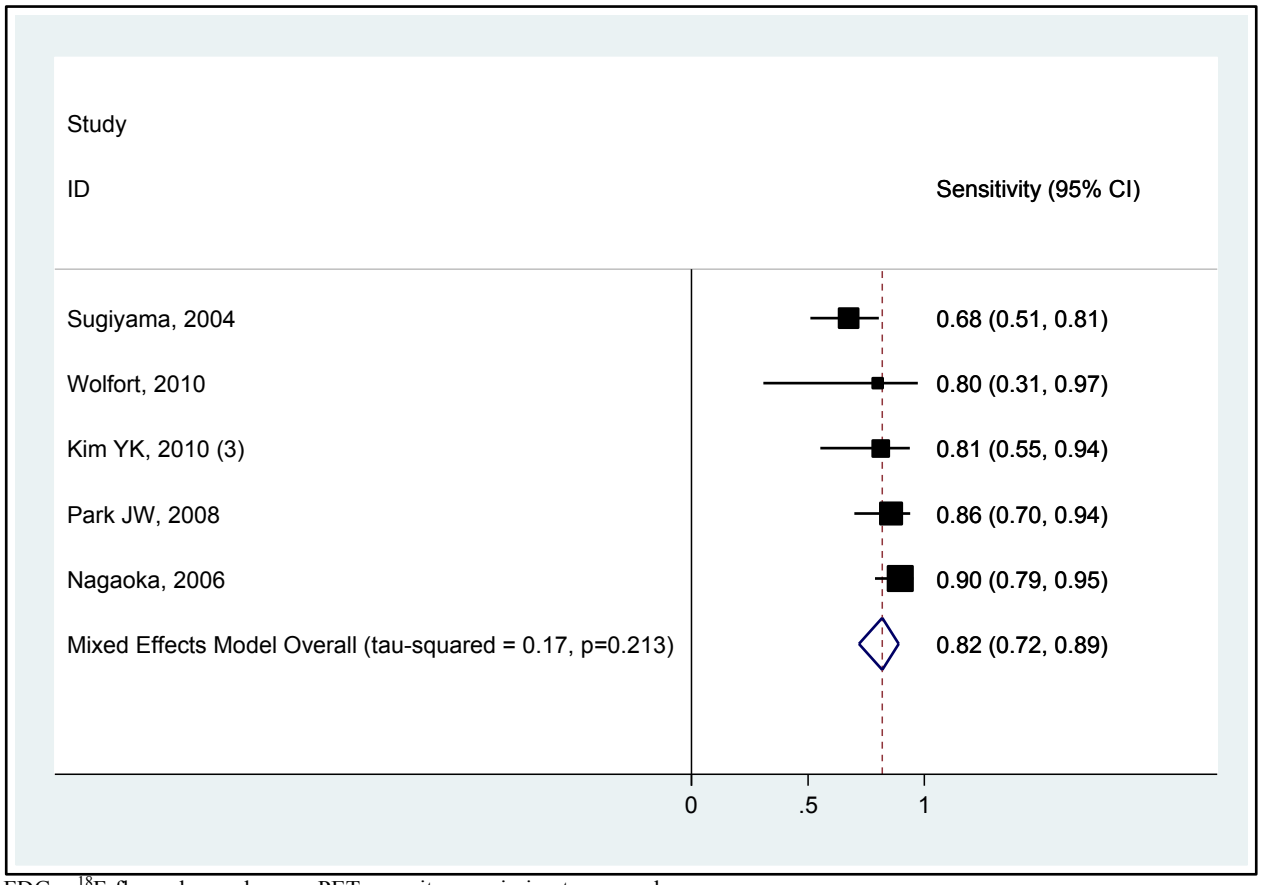
Figure 33. Test performance of FDG PET for detection of patients with metastatic hepatocellular carcinoma





FDG = ^{18}F-fluorodeoxyglucose; PET = positron emission tomography

Figure 34. Test performance of FDG PET for detection of metastatic hepatocellular carcinoma lesions



FDG = ^{18}F-fluorodeoxyglucose; PET = positron emission tomography

PET Versus Other Imaging Modalities

Three studies found no difference in sensitivity between PET/CT versus CT for metastatic HCC (0.82, 95% CI 0.61 to 0.93 vs. 0.85, 95% CI 0.66 to 0.95).[137,156,174]

One study found FDG PET associated with higher sensitivity than conventional imaging with CT, MRI, and chest x-ray for identifying HCC metastatic to lymph node (1.0 vs. 0.79) or bone (1.0 vs. 0.46), with no difference in specificity.[294] Both imaging methods identified all 12 patients with lung metastases. However, one other study found FDG PET associated with lower sensitivity than imaging with chest x-ray and CT for identifying lung metastases (1.0 vs. 0.61).[156]

KQ3.a.i. How is a particular technique's test performance modified by use of various reference standards?

Accuracy of Imaging for Staging

Evidence was insufficient to determine effects of the use of different reference standards on accuracy of imaging techniques for staging. All studies used explanted livers as the reference standard except for one, which used explanted livers or surgical resection as the reference standard.

PET for Detection of Metastatic Hepatocellular Carcinoma Disease

Evidence was too limited to determine effects of the use of different reference standards on accuracy of FDG PET for detection of metastatic HCC. Five of the six studies that used patients with metastatic HCC as the unit of analysis used mixed histological and imaging/clinical criteria as the reference standard. For studies that used metastatic HCC lesions as the reference standard, different reference standards (nonexplant histological reference standard, imaging and clinical criteria, or mixed) were each evaluated in only one or two studies.

KQ3.a.ii. How is the comparative effectiveness modified by patient, tumor, technical, or other factors?

Accuracy of Imaging for Staging

No study evaluated effects of patient-level characteristics or other factors on accuracy of imaging techniques for staging.

PET for Detection of Metastatic Hepatocellular Disease

Estimates of sensitivity stratified by use of PET or PET/CT showed no clear differences, with overlapping confidence intervals (Table 9). In one study that directly compared sensitivity of PET versus PET/CT for identifying metastatic HCC lesions, there was no clear difference in sensitivity (0.90 vs. 0.98, respectively, difference of -0.09, 95% CI -0.17 to 0.0) (Table 16).[174]

One study found PET with FDG associated with higher sensitivity than ^{11}C-acetate (0.79 vs. 0.64, for a difference of 0.15, 95% CI 0.03 to 0.28), but lower sensitivity than the combination of FDG and ^{11}C-acetate (0.98, difference -0.19, 95% CI -0.28 to -0.11).

Four studies of FDG PET that stratified analyses by metastatic lesion size found higher sensitivity as lesion size increased, but the number of lesions <1 cm was small (total of 20), precluding strong conclusions (Table 17).[156,279,283,285]

Eight studies reported test performance of FDG PET stratified by location of metastasis (Table 22).[137,156,174,279,283,285,291,294] In most studies, sensitivity was higher for lymph and bone metastasis than for lung metastasis, but samples were small, precluding strong conclusions.

KQ3.b. What is the comparative test performance of imaging techniques on diagnostic thinking?

Seven studies evaluated accuracy of imaging techniques for assessing transplant eligibility based on Milan criteria (Table 21).[75,92,109,114,164,187,192] Two studies also evaluated University of California, San Francisco (UCSF) liver transplantation criteria.[109,164] Six studies used explanted liver as the reference standard and the seventh[114] used an explanted liver or surgical resection. Five studies evaluated CT, one study evaluated MRI, one study evaluated PET, and two studies evaluated use of more than one imaging modality (US or CT and US, CT, or MRI). For CT, the proportion correctly assessed for transplant eligibility ranged from 40 percent to 96 percent using Milan criteria;[109,114,164,187,192] estimates from two studies that used UCSF criteria were similar to estimates based on Milan criteria. The proportion that met transplant criteria based on CT but exceeded criteria based on the reference standard was 3.5 to 7.8 percent in three studies[164,187,192] and 26 percent in the fourth.[109] In two studies, the proportion of patients who underwent transplantation based on CT but had no HCC lesion on examination of explanted livers was 2.3 percent and 16 percent.[109,187] In two studies not restricted to a single imaging modality (e.g., CT or MRI, or CT, MRI, or US), the proportion correctly assessed were 38 percent and 57 percent.[75,92]

Two studies directly compared accuracy of two or more imaging modalities for assessing transplant eligibility. One study reported similar accuracy for CT and MRI.[164] The other study found ^{11}C-choline PET more accurate than CT (95% vs. 40%).[114]

One study reported that the proportion of decisions to perform resection or ablative therapies that were classified as correct were similar for MRI (90% and 90%, respectively) and CT (80% and 77%, respectively).[113]

KQ3.c. What is the comparative effectiveness of imaging techniques on clinical and patient-centered outcomes?

One high risk of bias cohort study (n=167) of patients with HCC who underwent radiofrequency ablation compared effects of preprocedure US with contrast versus US without contrast plus CT on clinical decisionmaking (Appendix I).[44] US with contrast was performed within 10 minutes prior to the ablation procedure by two experienced radiologists using sulfur hexafluoride microbubble contrast; technical information for US without contrast was not reported. Contrast-enhanced CT with arterial and portal venous phase imaging was performed within one month prior to ablation, with followup one month after ablation. The study found that contrast enhanced US identified more small (≤2 cm) HCC lesions than US without contrast plus CT (36 vs. 31), and was associated with a higher complete necrosis rate following ablation (92% or 106/115 vs. 83% or 93/112 lesions, p=0.036). An important methodological shortcoming of this study was failure to adjust for potential confounders (Appendix C). Furthermore, additional lesions identified on US with contrast and cases classified as complete necrosis (treatment response) did not undergo histopathological or other confirmation prior to ablation.

KQ3.d. What are the adverse effects or harms associated with imaging-based staging strategies?

No study evaluated harms associated with use of imaging techniques for staging in patients diagnosed with HCC.

Table 6. Test performance of ultrasound imaging for identification and diagnosis of hepatocellular carcinoma

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
Surveillance (KQ 1), Ultrasound without contrast	Patient	0.82 (0.66 to 0.92); 0.34 (p=0.13)	3	0.87 (0.77 to 0.93); 0.17 (p=0.17)	2	6.2 (3.6 to 11)	0.20 (0.10-0.40)
	Lesion	0.60 (0.36 to 0.80); 0.34 (p=0.13)	2	0.94 (0.83 to 0.98)	1	9.8 (3.7 to 26)	0.43 (0.24 to 0.74)
Nonsurveillance (KQ 1), Ultrasound without contrast	Patient	0.73 (0.46 to 0.90); 2.3 (p=0.02)	8	0.93 (0.85 to 0.97); 0.78 (p=0.12)	6	11 (5.4 to 21)	0.29 (0.13 to 0.65)
• Excluding Doppler		0.77 (0.48 to 0.93); 2.5 (p=0.04)	7	0.92 (0.82 to 0.97); 0.70 (p=0.14)	5	9.8 (4.7 to 21)	0.25 (0.09 to 0.64)
• Prospective design		0.97 (0.68 to 0.998); 1.3 (p=0.02)	2	0.73 (0.45 to 0.90)	1	3.6 (1.5 to 8.5)	0.04 (0.003 to 0.59)
• Explant liver reference standard		0.48 (0.35 to 0.61); 0.17 (p=0.08)	5	0.96 (0.94 to 0.97); <0.0001 (p=0.97)	5	12 (7.4 to 19)	0.54 (0.42 to 0.70)
• Histopathological reference standard		0.95 (0.87 to 0.98)	3	0.73 (0.47 to 0.90)	1	3.6 (1.5 to 8.2)	0.07 (0.03 to 0.19)
• United States or Europe		0.70 (0.37 to 0.91); 1.9 (p=0.02)	5	0.93 (0.84 to 0.97); 0.51 (p=0.13)	4	10 (4.9 to 21)	0.32 (0.12 to 0.83)
• Excluding poor quality studies		0.77 (0.48 to 0.93); 2.5 (p=0.04)	7	0.92 (0.82 to 0.97); 2.5 (p=0.14)	5	9.8 (4.7 to 21)	0.25 (0.09 to 0.64)
• Avoided case-control design		0.54 (0.38 to 0.70); 0.44 (p=0.09)	6	0.95 (0.91 to 0.97); 0.41 (p=0.16)	6	11 (6.1 to 19)	0.48 (0.34 to 0.68
• Blinded interpretation of imaging		0.75 (0.33 to 0.95); 2.0 (p=0.02)	3	0.94 (0.81 to 0.98); 0.51 (p=0.13)	2	12 (4.4 to 33)	0.26 (0.07 to 0.97)
	Lesion	0.65 (0.49 to 0.78); 1.6 (p=0.0006)	16	0.83 (0.56 to 0.95); 0.74 (p=0.04)	2	3.8 (1.2 to 11)	0.43 (0.27 to 0.68)
• No contrast		0.60 (0.42 to 0.75); 1.3 (p=0.005)	11	0.63 (95% CI 0.53 to 0.73) and 0.95 (95% CI 0.85 to 0.99)	2	Not calculated	Not calculated
• With contrast		0.76 (0.53 to 0.90); 1.5 (p=0.03)	6	No data	--	--	--
• Perflubutane contrast		0.82 (0.64 to 0.92); 0.91 (p=0.07)	5	No data	--	--	--
• No Doppler		0.63 (0.45 to 0.78); 1.7 (p=0.0007)	14	Insufficient data	--	--	--
• Explanted liver		0.34 (0.21 to 0.49); 0.41 (p=0.007)	5	Insufficient data	--	--	--

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• Histopathological reference standard		0.70 (0.57 to 0.81)	7	Insufficient data	--	--	--
• Imaging and clinical criteria		0.75 (0.45 to 0.91)	1	Insufficient data	--	--	--
• Mixed histological and imaging/clinical criteria		0.85 (0.72 to 0.93)	3	Insufficient data	--	--	--
• Prospective		0.72 (0.46 to 0.89); 1.6 (p=0.0009)	6	Insufficient data	--	--	--
• United States or Europe		0.57 (0.27 to 0.83); 1.7 (p=0.0007)	5	Insufficient data	--	--	--
• Excluding poor quality studies		0.58 (0.41 to 0.73); 1.4 (p=0.001)	13	Insufficient data	--	--	--
• Avoided case-control design		0.43 (0.21 to 0.68); 1.3 (p=0.001)	5	Insufficient data	--	--	--
• Blinded interpretation of imaging		0.73 (0.55 to 0.85); 1.3 (p=0.001)	10	Insufficient data	--	--	--
	Liver segment	0.79 (0.62 to 0.89); 1.6 (p=0.0006)	2	0.95 (0.84 to 0.99); 0.74 (p=0.04)	2	17 (4.7 to 60)	0.22 (0.12 to 0.42)
Evaluation of previously identified lesion (KQ 2)	Patient	0.87 (0.78 to 0.93); 0.71 (p=0.01)	9	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	10 (5.4 to 20)	0.14 (0.08 to 0.24)
• No contrast		0.78 (0.72 to 0.83); <0.0001 (p=1.0)	2	No data	--	--	--
• With contrast		0.88 (0.79 to 0.94); 0.78 (p=0.02)	8	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	11 (5.5 to 20)	0.13 (0.07 to 0.24)
• Sulfur hexafluoride contrast		0.82 (0.69 to 0.90); 0.43 (p=0.02)	5	0.92 (0.81 to 0.97); 0.35 (p=0.09)	3	10 (4.1 to 24)	0.20 (0.11 to 0.35)
• Perflubutane contrast		0.94 (0.87 to 0.98)	3	0.92 (0.78 to 0.97)	2	11 (4.0-31)	0.06 (0.03 to 0.15)
• Excluding Doppler		0.87 (0.77 to 0.93); 0.80 (p=0.02)	8	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	10 (5.4 to 20)	0.14 (0.08 to 0.26)
• With Doppler		0.64 (0.39 to 0.83); 1.5 (p=0.03)	6	0.93 (0.87 to 0.97); 0.02 (p=0.75)	3	9.6 (4.6 to 20)	0.38 (0.20-0.84)
• Histopathological reference standard		0.83 (0.64 to 0.93); 0.66 (p=0.02)	3	0.91 (0.77 to 0.97); 0.35 (p=0.09)	2	9.4 (3.3 to 26)	0.19 (0.08 to 0.44)
• Imaging and clinical criteria		0.78 (0.69 to 0.86)	1	Insufficient data			
• Mixed histological and imaging/clinical criteria		0.91 (0.81 to 0.96)	5	0.92 (0.82 to 0.97)	3	11 (4.9 to 26)	0.10 (0.05 to 0.22)
• Prospective Design		0.84 (0.74 to 0.90); 0.49 (p=0.01)	7	0.92 (0.81 to 0.97); 0.35 (p=0.09)	3	10 (4.1 to 24)	0.20 (0.13 to 0.31)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• Excluding studies restricted to hypervascular HCC		0.87 (0.77 to 0.93); 0.80 (p=0.02)	8	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	10 (5.4 to 20)	0.14 (0.08 to 0.26)
• United States or Europe		0.82 (0.68 to 0.90); 0.55 (p=0.01)	5	0.92 (0.81 to 0.97); 0.35 (p=0.09)	3	10 (4.1 to 25)	0.20 (0.11 to 0.37)
• Excluding poor quality studies		0.84 (0.72 to 0.91); 0.55 (p=0.01)	6	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	9.9 (4.1 to 19)	0.18 (0.10-0.32)
• Avoided case-control design		0.88 (0.78 to 0.94); 0.71 (p=0.01)	7	0.92 (0.84 to 0.96); 0.35 (p=0.09)	5	10 (5.4 to 20)	0.13 (0.07 to 0.25)
• Blinded interpretation of imaging		0.93 (0.87 to 0.96); 0.30 (p=0.02)	5	0.94 (0.89 to 0.97); 0.08 (p=0.51)	3	15 (8.0-28)	0.08 (0.04 to 0.15)
	Lesion	0.82 (0.72 to 0.89); 1.8 (p<0.0001)	22	0.91 (0.87 to 0.95); 0.41 (p=0.005)	13	9.4 (6.2 to 14)	0.19 (0.12 to 0.31)
• No contrast		0.62 (0.18 to 0.93); 3.7 (p=0.19)	4	0.92 (0.84 to 0.96) (p=0.34)	3	8.1 (3.6 to 18)	0.41 (0.12 to 1.4)
• With contrast		0.86 (0.79 to 0.91); 1.0 (p<0.0001)	21	0.93 (0.87 to 0.96); 0.64 (p=0.01)	11	12 (6.3 to 21)	0.15 (0.10-0.23)
• Sulfur hexafluoride contrast		0.86 (0.77 to 0.92); 1.0 (p=0.003)	12	0.92 (0.75 to 0.98); 2.6 (p=0.09)	6	11 (3.3 to 36)	0.15 (0.09 to 0.25)
• Perflubutane contrast		0.82 (0.51 to 0.95); 1.6 (p=0.26)	3	0.92 (0.86 to 0.96); 0.07 (p=0.52)	2	10 (4.8 to 21)	0.20 (0.06 to 0.67)
• Galactose contrast		0.90 (0.86 to 0.94); <0.0001 (p=1.0)	4	0.95 (0.86 to 0.98); <0.0001 (p=1.0)	2	18 (6.0-55)	0.10 (0.07 to 0.15)
• Imaging and clinical criteria		0.89 (0.75 to 0.95); 1.7 (p<0.0001)	9	0.88 (0.79 to 0.93); 0.26 (p=0.02)	6	7.4 (4.2 to 13)	0.13 (0.06 to 0.29)
• Mixed histological and imaging/clinical criteria		0.77 (0.62 to 0.87)	15	0.93 (0.88 to 0.96)	8	11 (6.5 to 19)	0.25 (0.14 to 0.43)
• Prospective design		0.87 (0.74 to 0.94); 1.6 (p<0.0001)	9	0.93 (0.87 to 0.96); 0.26 (p=0.03)	5	12 (6.5 to 23)	0.15 (0.07 to 0.30)
• Excluding studies restricted to hypervascular HCC		0.82 (0.71 to 0.89); 1.9 (p<0.0001)	21	0.91 (0.87 to 0.94); 0.42 (p=0.005)	12	9.4 (6.2 to 14)	0.20 (0.12 to 0.33)
• United States or Europe		0.82 (0.65 to 0.92); 1.8 (p<0.0001)	10	0.87 (0.75 to 0.93); 0.31 (p=0.01)	4	6.1 (3.2 to 12)	0.21 (0.10-0.43)
• No Doppler		0.85 (0.76 to 0.92); 1.6 (p<0.0001)	19	0.91 (0.85 to 0.95); 0.53 (p=0.01)	10	9.4 (5.5 to 16)	0.16 (0.10-0.27)
• Excluding poor quality studies		0.80 (0.67 to 0.89); 1.8 (p<0.0001)	16	0.92 (0.87 to 0.95); 0.40 (p=0.005)	11	9.6 (6.1 to 15)	0.22 (0.13 to 0.37)
• Avoided case-control design		0.79 (0.66 to 0.88); 1.6 (p<0.0001)	17	0.91 (0.86 to 0.94); 0.37 (p=0.006)	11	8.6 (5.7 to 13)	0.23 (0.14 to 0.38)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• Blinded interpretation of imaging		0.81 (0.67 to 0.89); 1.8 (p<0.0001)	15	0.90 (0.85 to 0.94); 0.34 (p=0.007)	9	8.1 (5.4 to 12)	0.22 (0.13 to 0.38)
• Used confidence rating scale		0.62 (0.31 to 0.85); 1.6 (p<0.0001)	3	0.89 (0.79 to 0.95); 0.35 (p=0.0087)	3	5.6 (2.9 to 11)	0.43 (0.20-0.91)
Lesion depth <53, <60, or <85 mm		0.87 (0.80-0.92); <0.0001 (p=1.0)	3	Insufficient data	--	--	--
• >53, >60, or >85 mm		0.83 (0.74 to 0.90)	3	Insufficient data	--	--	--
Body mass index <23 or <25		0.80 (0.70-0.88); 0.11 (p=0.37)	2	Insufficient data	--	--	--
• >23 or >25		0.80 (0.70-0.87)	2	Insufficient data	--	--	--

HCC = hepatocellular carcinoma; KQ = Key Question

Table 7. Test performance of computed tomography imaging for identification of intrahepatic and extrahepatic hepatocellular carcinoma

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
Surveillance	Patient	0.84 (0.59 to 0.95); 0.50 (p=0.74)	2	0.99 (0.86 to 0.999); 1.1 (p=0.97)	2	60 (5.9 to 622)	0.16 (0.06 to 0.47)
	Lesion	0.62 (0.46 to 0.76)	1	Insufficient data	--	--	--
Nonsurveillance	Patient	0.83 (0.75 to 0.89); 0.60 (p=0.008)	16	0.92 (0.86 to 0.96); 0.63 (p=0.02)	11	11 (5.6 to 20)	0.19 (0.12 to 0.28)
• Excluding studies restricted to hypervascular HCC		0.84 (0.77 to 0.90); 0.46 (p=0.03)	15	0.92 (0.85 to 0.96); 0.62 (p=0.02)	11	11 (5.6 to 21)	0.17 (0.11 to 0.26)
• Multidetector CT, ≥8 rows		0.88 (0.69 to 0.96); 0.55 (p=0.01)	2	Insufficient data	--	--	--
• Multidetector CT, <8 rows		0.89 (0.69 to 0.97)	2	0.98 (0.87 to 0.996)	1	37 (6.1 to 222)	0.11 (0.03 to 0.37)
• Non-multidetector CT		0.82 (0.71 to 0.89)	11	0.90 (0.83 to 0.95); 0.48 (p=0.07)	9	8.3 (4.5 to 15)	0.20 (0.12 to 0.34)
• Contrast rate ≥3 ml/s		0.87 (0.77 to 0.93); 0.48 (p=0.04)	8	0.90 (0.81 to 0.94); 0.38 (p=0.08)	7	8.4 (4.5 to 16)	0.15 (0.08 to 0.27)
• Contrast rate <3 ml/s		0.71 (0.50-0.85)	4	0.95 (0.87 to 0.98)	3	13 (4.8 to 37)	0.31 (0.17 to 0.58)
• Arterial, portal venous, and delayed phase imaging		0.89 (0.81 to 0.94); 0.33 (p=0.06)	7	0.88 (0.74 to 0.95); 0.41 (p=0.09)	4	7.3 (3.2 to 17)	0.13 (0.07 to 0.22)
• Missing delayed phase imaging		0.78 (0.66 to 0.87)	7	0.94 (0.87 to 0.97)	5	13 (5.8 to 27)	0.23 (0.14 to 0.38)
• Delayed phase ≥120 s		0.87 (0.79 to 0.92); 0.17 (p=0.12)	6	0.87 (0.74 to 0.94); 0.40 (p=0.10)	4	6.8 (3.1 to 15)	0.15 (0.09 to 0.25)
• Section thickness ≤5 mm		0.84 (0.73 to 0.91); 0.57 (p=0.02)	9	0.87 (0.76 to 0.93); 0.19 (p=0.19)	6	6.3 (3.4 to 12)	0.19 (0.11 to 0.33)
• Section thickness >5 mm		0.83 (0.67 to 0.92)	5	0.95 (0.90-0.98)	4	17 (8.3 to 34)	0.18 (0.09 to 0.37)
• Explanted liver reference standard		0.81 (0.71 to 0.88); 0.47	11	0.93 (0.87 to 0.96); 0.36	9	11 (6.0-20)	0.21 (0.13 to 0.32)
• Other histopathological reference standard		0.85 (0.64 to 0.95)	3	0.73 (0.31 to 0.94)	1	3.1 (0.83 to 12)	0.20 (0.07 to 0.62)
• Histological and clinical/imaging reference standard		0.88 (0.70-0.96)	2	0.93 (0.75 to 0.98)	1	13 (3.2 to 52)	0.13 (0.05 to 0.35)
• Prospective		0.72 (0.57 to 0.84); 0.30 (p=0.03)	5	0.85 (0.67 to 0.94); 0.43 (p=0.03)	3	4.9 (2.0-12)	0.33 (0.20-0.54)
• United States or Europe		0.83 (0.73 to 0.90); 0.60 (p=0.008)	12	0.89 (0.80-0.94); 0.32 (p=0.11)	8	7.3 (3.9 to 14)	0.19 (0.12 to 0.32)
• Used confidence rating scale		0.85 (0.70-0.93); 0.53 (p=0.01)	4	0.73 (0.31 to 0.94); 0.36 (p=0.05)	1	3.1 (0.84 to 12)	0.20 (0.08 to 0.51)
• Excluding poor quality studies		0.85 (0.75 to 0.91); 0.55 (p=0.01)	9	0.92 (0.83 to 0.97); 0.63 (p=0.02)	6	11 (4.6 to 26)	0.16 (0.10-0.28)
• Avoided case-control design		0.75 (0.63 to 0.84); 0.44 (p=0.007)	9	0.91 (0.83 to 0.95); 0.59 (p=0.03)	9	8.1 (4.1 to 16)	0.28 (0.18 to 0.43)
• Blinded interpretation of		0.83 (0.74 to 0.0.89);	13	0.94 (0.88 to 0.97); 0.49	8	13 (6.5 to 27)	0.19 (0.12 to 0.29)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
imaging		0.55 (p=0.01)		(p=0.02)			
	Lesion	0.77 (0.73 to 0.81); 0.82 (p<0.0001)	75	0.89 (0.83 to 0.93); 0.93 (p<0.0001)	20	7.0 (4.6 to 11)	0.25 (0.21 to 0.30)
• Multidetector CT, ≥8 rows		0.76 (0.70-0.81); 0.81 (p<0.0001)	37	0.92 (0.85 to 0.96); 0.83 (p=<0.0001)	5	9.4 (5.0-17)	0.26 (0.20-0.33)
• Multidetector CT, <8 rows		0.79 (0.59 to 0.91)	4	Insufficient data	--	--	--
• Non-multidetector CT		0.76 (0.69 to 0.83)	23	0.85 (0.73 to 0.92)	8	5.0 (2.7 to 9.3)	0.28 (0.20-0.38)
• Contrast rate ≥3 ml/s		0.79 (0.75 to 0.83); 0.79 (p<0.0001)	62	0.88 (0.82 to 0.93); 0.88 (p<0.0001)	18	6.8 (4.4 to 11)	0.24 (0.20-0.29)
• Contrast rate <3 ml/s		0.72 (0.58 to 0.83)	8	0.93 (0.75 to 0.99)	2	11 (2.6 to 47)	0.29 (0.18 to 0.47)
• Arterial, portal venous, and delayed phase imaging		0.75 (0.69 to 0.80); 0.79 (p<0.0001)	44	0.91 (0.86 to 0.95); 0.62 (p<0.0001)	14	8.4 (5.2 to 14)	0.28 (0.22 to 0.34)
• Missing delayed phase imaging		0.81 (0.75 to 0.86); 0.78 (p<0.0001)	27	0.87 (0.73 to 0.94); 0.49 (p<0.0001)	5	6.2 (1.3 to 11)	0.22 (0.16 to 0.30)
• Delayed phase ≥120 s		0.74 (0.68 to 0.80); 0.78 (p<0.0001)	39	0.93 (0.88 to 0.96); 0.49 (p<0.0001)	11	10 (6.1 to 17)	0.28 (0.22 to 0.34
• Section thickness ≤5 mm		0.75 (0.71 to 0.80); 0.79 (p<0.0001)	56	0.90 (0.85 to 0.94); 0.85 (p<0.0001)	17	7.6 (4.8 to 12)	0.27 (0.22 to 0.33)
• Section thickness >5 mm		0.83 (0.74 to 0.89)	14	0.81 (0.48 to 0.95)	2	4.4 (1.2 to 16)	0.21 (0.13 to 0.35)
• Excluding studies restricted to hypervascular HCC		0.76 (0.71 to 0.81); 0.82 (p<0.0001)	56	0.87 (0.81 to 0.92); 0.77 (p<0.0001)	17	5.9 (3.9 to 9.0)	0.27 (0.22 to 0.33)
• Excluding studies restricted to HCC lesions <2 cm		0.78 (0.73 to 0.81); 0.82 (p<0.0001)	69	0.88 (0.82 to 0.93); 0.89 (p<0.0001)	18	6.6 (4.2 to 10)	0.25 (0.21 to 0.31)
• Explanted liver reference standard		0.69 (0.60-0.77); 0.70 (p<0.0001)	21	0.81 (0.72 to 0.88); 0.54 (p<0.0001)	11	3.6 (2.4 to 5.4)	0.38 (0.30-0.50)
• Other histopathological reference standard		0.85 (0.77 to 0.90)	13	0.95 (0.89 to 0.97)	5	16 (7.7 to 32)	0.16 (0.11 to 0.25)
• Imaging/clinical reference standard		0.65 (0.41 to 0.83)	3	Insufficient data	--	--	--
• Mixed histological and imaging/clinical reference standard		0.79 (0.74 to 0.84)	38	0.92 (0.85 to 0.96)	4	10 (5.1 to 21)	0.22 (0.18 to 0.28)
• Prospective		0.74 (0.64 to 0.81); 0.71 (p<0.0001)	17	0.84 (0.73 to 0.91); 0.70 (p<0.0001)	8	4.6 (2.6 to 8.1)	0.31 (0.22 to 0.44)
• United States or Europe		0.76 (0.68 to 0.83); 0.81 (p<0.0001)	23	0.79 (0.67 to 0.87); 0.61 (p<0.0001)	10	3.6 (2.2 to 5.7)	0.30 (0.22 to 0.41)
• Used confidence rating scale		0.75 (0.69 to 0.80); 0.80 (p<0.0001)	35	0.93 (0.88 to 0.96); 0.68 (p<0.0001)	11	10 (5.9 to 17)	0.27 (0.21 to 0.34)
• Excluding poor quality studies		0.75 (0.70-0.80); 0.79 (p<0.0001)	48	0.91 (0.87 to 0.94); 0.61 (p<0.0001)	16	8.6 (5.7 to 13)	0.27 (0.22 to 0.34)
• Avoided case-control design		0.74 (0.66 to 0.81); 0.80 (p<0.0001)	21	0.87 (0.77 to 0.93); 0.87 (p<0.0001)	10	5.7 (3.1 to 10)	0.30 (0.21 to 0.41)
• Blinded interpretation of imaging		0.77 (0.71 to 0.81); 0.82 (p<0.0001)	45	0.88 (0.80-0.93); 0.93 (p<0.0001)	11	6.6 (3.7 to 12)	0.24 (0.18 to 0.32)
	Liver	0.90 (0.87 to 0.93);	7	0.97 (0.94 to 0.98); 0.34	7	26 (15 to 45)	0.10 (0.07 to 0.13)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
	segment	0.0001 (p=0.95)		(p=0.07)			
• Excluding studies restricted to hypervascular HCC		0.90 (0.87 to 0.93); 0.004 (p=0.76)	6	0.97 (0.94 to 0.98); 0.37 (p=0.10)	6	30 (16 to 57)	0.10 (0.08 to 0.14)
• Excluding studies restricted to HCC lesions <2 cm		0.90 (0.87 to 0.93); 0.0001 (p=0.96)	6	0.97 (0.94 to 0.98)	6	29 (15 to 56)	0.10 (0.07 to 0.13)
Evaluation of previously identified lesion	Patient	0.85 (0.76 to 0.91); 0.28 (p=0.12)	5	0.92 (0.86 to 0.96); 0.11 (p=0.58)	3	11 (5.7 to 22)	0.17 (0.10-0.27)
	HCC lesion	0.80 (0.67 to 0.88); 1.1 (p=0.002)	12	0.89 (0.29 to 0.99); 4.2 (p=0.14)	6	6.9 (0.53—91)	0.23 (0.13 to 0.40)
• Multidetector CT, ≥8 rows		0.57 (0.37 to 0.74); 0.40 (p=0.007)	3	0.96 (0.67 to 0.996); 3.4 (p=0.06)	3	14 (1.7 to 113)	0.45 (0.31 to 0.66)
• Multidetector CT, <8 rows		0.62 (0.31 to 0.85)	2	0.69 (0.05 to 0.99)	2	2.0 (0.25 to 16)	0.55 (0.35 to 0.87)
• Non-multidetector CT		0.86 (0.77 to 0.92)	6	0.96 (0.50-0.998)	1	20 (1.1 to 378)	0.14 (0.09 to 0.23)
• Contrast rate ≥3 ml/s		0.80 (0.67 to 0.88); 1.1 (p=0.002)	12	0.89 (0.29 to 0.99); 4.2 (p=0.14)	6	6.9 (0.53—91)	0.23 (0.13 to 0.40)
• Arterial, portal venous, and delayed phase imaging		0.75 (0.62 to 0.84); 0.74 (p=0.002)	10	0.91 (0.50-0.99); 3.8 (p=0.10)	6	8.4 (1.0-67)	0.28 (0.18 to 0.44)
• Missing delayed phase imaging		0.94 (0.79 to 0.98)	2	Insufficient data	--	--	--
• Delayed phase imaging >120 s		0.75 (0.61 to 0.86); 0.77 (p=0.003)	8	0.48 (0.04 to 0.96); 8.8 (p=0.09)	5	1.5 (0.34 to 6.2)	0.51 (0.13 to 2.0)
• Section thickness ≤5 mm		0.69 (0.55 to 0.80); 0.52 (p=0.003)	8	0.93 (0.66 to 0.99); 3.6 (p=0.08)	6	10 (1.7 to 58)	0.33 (0.23 to 0.49)
• Section thickness >5 mm		0.91 (0.77 to 0.97)	2	Insufficient data	--	--	--
• Excluding studies restricted to HCC lesions <2 cm		0.87 (0.78 to 0.93); 0.58 (p=0.004)	7	0.63 (0.03 to 0.99); 2.2 (p=0.08)	1	2.3 (0.20-28)	0.21 (0.04 to 0.96)
• Histopathological reference standard		0.77 (0.62 to 0.88); 1.0 (p=0.003)	9	0.89 (0.50-0.99); 2.1 (p=0.13)	5	7.1 (1.1 to 46)	0.26 (0.14 to 0.46)
• Histological and clinical/imaging reference standard		0.85 (0.63 to 0.95)	3	0.98 (0.14 to 1.0)	1	46 (0.16 to 12900)	0.15 (0.05 to 0.42)
• Prospective		0.81 (0.62 to 0.91); 1.0 (p=0.002)	6	0.83 (0.56 to 0.95); 0.48 (p=0.20)	3	4.8 (1.6 to 15)	0.23 (0.11 to 0.50)
• United States or Europe		0.79 (0.63 to 0.89); 1.1 (p=0.002)	9	0.33 (0.01 to 0.96); 12 (p=0.09)	5	1.2 (0.37 to 3.8)	0.63 (0.07 to 5.7)
• Used confidence rating scale		0.86 (0.58 to 0.96); 1.0 (p=0.003)	2	0.67 (0.05 to 0.99)	1	2.6 (0.22 to 31)	0.21 (0.04 to 1.1)
• Excluding poor quality studies		0.71 (0.54 to 0.83); 0.79 (p=0.003)	8	0.66 (0.11 to 0.97); 9.0 (p=0.09)	5	2.1 (0.39 to 11)	0.45 (0.23 to 0.87)
• Avoided case-control design		0.88 (0.72 to 0.95); 0.90 (p=0.003)	4	Insufficient data	--	--	--
• Blinded interpretation of imaging		0.80 (0.51 to 0.94); 0.91 (p=0.005)	3	0.50 (0.01 to 0.99); 2.1 (p=0.22)	2	1.6 (0.21 to 12)	0.40 (0.05 to 3.2

CT = computed tomography; HCC = hepatocellular carcinoma

Table 8. Test performance of magnetic resonance imaging for identification of intrahepatic and extrahepatic hepatocellular carcinoma

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
Nonsurveillance	Patient	0.87 (0.77 to 0.93); 0.70 (p=0.05)	11	0.88 (0.79 to 0.93); 0.58 (p=0.04)	9	7.2 (3.9 to 13)	0.15 (0.08 to 0.27)
• Excluding studies restricted to hypervascular HCC		0.87 (0.76 to 0.93); 0.87 (p=0.05)	10	0.88 (0.79 to 0.94); 0.70 (p=0.05)	8	7.5 (3.8 to 15)	0.15 (0.08 to 0.29)
• 1.5 T MRI		0.85 (0.76 to 0.91); 0.45 (p=0.10)	10	0.90 (0.82 to 0.95); 0.48 (p=0.08)	8	8.3 (4.5 to 15)	0.17 (0.10-0.27)
• Gadopentetate or gadodiamide contrast		0.85 (0.76 to 0.91); 0.45 (p=0.10)	10	0.90 (0.82 to 0.95); 0.48 (p=0.08)	8	8.3 (4.5 to 15)	0.17 (0.10-0.27)
• Arterial, portal venous, and delayed phase imaging		0.87 (0.77 to 0.93); 0.68 (p=0.05)	9	0.87 (0.76 to 0.93); 0.64 (p=0.04)	7	6.6 (3.4 to 13)	0.15 (0.08 to 0.29)
• Missing delayed phase imaging		0.84 (0.53 to 0.96)	2	0.92 (0.68 to 0.99)	2	11 (2.0-62)	0.18 (0.05 to 0.67)
• Delayed phase ≥120 s		0.78 (0.52 to 0.92); 0.31 (p=0.15)	2	0.92 (0.77 to 0.97); 0.40 (p=0.10)	2	9.6 (3.0-30)	0.24 (0.09 to 0.60)
• Section thickness ≤5 mm		0.88 (0.74 to 0.95); 0.56 (p=0.06)	4	0.94 (0.85 to 0.98); 0.16 (p=0.40)	3	16 (5.8 to 42)	0.12 (0.05 to 0.30)
• Section thickness >5 mm		0.76 (0.50-0.91)	3	0.85 (0.68 to 0.94)	3	5.0 (2.0-12)	0.29 (0.12 to 0.71)
• Explanted liver reference standard		0.89 (0.79 to 0.94); 0.67 (p=0.10)	9	0.90 (0.81 to 0.95); 0.63 (p=0.08)	7	9.2 (4.4 to 19)	0.12 (0.07 to 0.24)
• Prospective		0.77 (0.55 to 0.91); 0.68 (p=0.07)	4	0.94 (0.84 to 0.98); 0.30 (p=0.16)	3	12 (4.4 to 34)	0.24 (0.11 to 0.54)
• United States or Europe		0.87 (0.77 to 0.93); 0.70 (p=0.05)	11	0.88 (0.79 to 0.93); 0.58 (p=0.04)	9	7.2 (3.9 to 13)	0.15 (0.08 to 0.27)
• Used confidence rating scale		0.87 (0.55 to 0.97); 0.72 (p=0.06)	2	0.72 (0.50-0.87); 0.22 (p=0.14)	2	3.1 (1.4 to 7.0)	0.19 (0.04 to 0.87)
• Excluding poor quality studies		0.85 (0.75 to 0.92); 0.66 (p=0.06)	9	0.87 (0.77 to 0.93); 0.57 (p=0.04)	7	6.6 (3.4 to 13)	0.17 (0.09 to 0.31)
• Avoided case-control design		0.84 (0.74 to 0.91); 0.49 (p=0.08)	9	0.87 (0.77 to 0.93); 0.49 (p=0.05)	8	6.4 (3.5 to 12)	0.18 (0.11 to 0.32)
• Blinded interpretation of imaging		0.91 (0.82 to 0.96); 0.47 (p=0.08)	5	0.94 (0.88 to 0.97); 0.14 (p=0.16)	4	15 (7.2 to 31)	0.09 (0.04 to 0.20)
	Lesion	0.83 (0.80-0.86); 0.76 (p<0.0001)	69	0.83 (0.70-0.92); 1.8 (p=<0.0001)	13	5.0 (2.6 to 9.6)	0.20 (0.16 to 0.26)
• 3.0 T MRI		0.89 (0.81 to 0.94); 0.73 (p<0.0001)	9	0.96 (0.60-0.998); 1.5 (p=<0.0001)	1	25 (1.5 to 408)	0.12 (0.07 to 0.21)
• 1.5 T MRI		0.82 (0.78 to 0.85)	57	0.81 (0.66 to 0.91)	11	4.4 (2.3 to 8.4)	0.22 (0.17 to 0.29)
• Gadopentetate or gadodiamide contrast		0.80 (0.75 to 0.84); 0.69 (p<0.0001)	38	0.63 (0.49 to 0.75); 0.53 (p<0.0001)	6	2.1 (1.5 to 3.0)	0.32 (0.25 to 0.42)
• Gadobenate disodium or gadobenate contrast		0.86 (0.81 to 0.90)	30	0.96 (0.93 to 0.98)	7	22 (12 to 41)	0.15 (0.11 to 0.19)
• Arterial, portal venous, and delayed phase imaging		0.83 (0.79 to 0.86); 0.74 (p<0.0001)	65	0.80 (0.64 to 0.90); 2.0 (p=0.0002)	12	4.2 (2.1 to 8.2)	0.21 (0.16 to 0.28)
• Missing delayed phase imaging		0.77 (0.52 to 0.91)	3	0.94 (0.50-0.996)	1	14 (0.87 to 218)	0.24 (0.09 to 0.63)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• Delayed phase ≥120 s		0.84 (0.80-0.87); 0.72 (p<0.0001)	56	0.84 (0.65 to 0.93); 1.9 (p=0.004)	9	5.1 (2.2 to 12)	0.20 (0.14 to 0.27)
• Section thickness ≤5 mm		0.87 (0.84 to 0.90); 0.55 (p<0.0001)	41	0.83 (0.67 to 0.92); 1.8 (p<0.0001)	8	5.1 (2.4 to 11)	0.16 (0.12 to 0.21)
• Section thickness >5 mm		0.81 (0.58 to 0.93)	17	0.87 (0.60-0.97)	3	6.2 (1.6 to 23)	0.23 (0.16 to 0.35)
• Excluding studies that used diffusion-weighted imaging		0.83 (0.79 to 0.86); 0.78 (p<0.0001)	66	0.82 (0.68 to 0.91); 1.6 (p=0.0002)	12	4.7 (2.4 to 8.9)	0.21 (0.16 to 0.27)
• Excluding studies restricted to hypervascular HCC		0.85 (0.81 to 0.87); 0.65 (p<0.0001)	57	0.84 (0.72 to 0.92); 1.8 (p<0.0001)	13	5.4 (2.8 to 10)	0.18 (0.14 to 0.23)
• Excluding studies restricted to HCC lesion <2 cm		0.83 (0.79 to 0.86); 0.77 (p<0.0001)	57	0.83 (0.68 to 0.92); 1.8 (p=0.0002)	12	4.8 (2.4 to 9.5)	0.21 (0.16 to 0.27)
• Explanted liver reference standard		0.69 (0.59 to 0.77); 0.58 (p<0.0001)	15	0.78 (0.57 to 0.91); 0.96 (p=0.002)	6	3.2 (1.5 to 7.0)	0.39 (0.28 to 0.55)
• Other histopathological reference standard		0.88 (0.81 to 0.93)	11	0.97 (0.88 to 0.99)	4	27 (7.2 to 98)	0.12 (0.08 to 0.20)
• Imaging/clinical reference standard		0.86 (0.66 to 0.95)	2	Insufficient data	--	--	--
• Mixed histological and imaging/clinical reference standard		0.85 (0.82 to 0.88)	41	0.81 (0.49 to 0.95)	3	4.4 (1.4 to 14)	0.18 (0.13 to 0.26)
• Prospective		0.84 (0.76 to 0.89); 0.76 (p<0.0001)	19	0.89 (0.72 to 0.96); 1.8 (p<0.0001)	5	7.8 (2.7 to 23)	0.18 (0.12 to 0.28)
• United States or Europe		0.77 (0.69 to 0.84); 0.71 (p<0.0001)	19	0.71 (0.55 to 0.83); 0.70 (p=0.002)	7	2.7 (1.6 to 4.4)	0.32 (0.23 to 0.46)
• Used confidence rating scale		0.87 (0.83 to 0.89); 0.61 (p<0.0001)	43	0.88 (0.75 to 0.95); 1.5 (p=0.0001)	9	7.5 (3.2 to 18)	0.15 (0.12 to 0.20)
• Excluding poor quality studies		0.82 (0.78 to 0.86); 0.74 (p<0.0001)	48	0.84 (0.68 to 0.93); 1.7 (p<0.0001)	16	5.1 (2.4 to 11)	0.22 (0.16 to 0.29)
• Avoided case-control design		0.81 (0.74 to 0.87); 0.77 (p<0.0001)	21	0.78 (0.61 to 0.89); 1.2 (p=0.0002)	10	3.7 (1.9 to 7.0)	0.24 (0.16 to 0.36)
• Blinded interpretation of imaging		0.84 (0.80-0.87); 0.75 (p<0.0001)	44	0.86 (0.69 to 0.94); 1.7 (p<0.0001)	8	5.9 (2.5 to 14)	0.19 (0.14 to 0.25)
Evaluation of previously identified lesion	Patient	0.76 (0.62 to 0.86); 0.20 (p=0.35)	3	0.87 (0.70-0.95); 0.51 (p=0.38)	3	5.9 (2.5 to 14)	0.28 (0.18 to 0.43)
	HCC lesion	0.79 (0.69 to 0.87); 0.70 (p=0.004)	13	0.95 (0.82 to 0.99); 3.4 (p=0.01)	12	15 (4.4 to 50)	0.22 (0.15 to 0.33)
• 3.0 T MRI		0.89 (0.64 to 0.97); 0.63 (p=0.004)	2	0.82 (0.13 to 0.994); 3.1 (p=0.01)	1	4.9 (0.31 to 75)	0.14 (0.04 to 0.49)
• 1.5 T MRI		0.77 (0.66 to 0.86)	11	0.96 (0.85 to 0.99)	11	18 (5.0-67)	0.24 (0.16 to 0.36)
• Gadopentetate or gadodiamide contrast		0.82 (0.65 to 0.92); 0.75 (p=0.004)	4	0.95 (0.85 to 0.99); 0.87 (p=0.02)	4	18 (5.8 to 57)	0.19 (0.09 to 0.38)
• Gadoxetic acid or gadobenate contrast		0.77 (0.61 to 0.88)	7	0.95 (0.88 to 0.98)	6	17 (7.3 to 38)	0.24 (0.13 to 0.42)
• Arterial, portal venous, and delayed phase imaging		0.78 (0.67 to 0.86); 0.63 (p=0.005)	12	0.95 (0.82 to 0.99); 3.3 (p=0.01)	11	16 (4.4 to 57)	0.23 (0.16 to 0.35)
• Missing delayed phase imaging		0.90 (0.60-0.98)	1	0.89 (0.16 to 0.997)	1	8.4 (0.31 to 227)	0.11 (0.03 to 0.48)

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• Delayed phase imaging >120 s		0.79 (0.67 to 0.87); 0.59 (p=0.006)	10	0.96 (0.80-0.992); 3.9 (p=0.02)	10	18 (3.8 to 84)	0.22 (0.15 to 0.34)
• Section thickness <=5 mm		0.69 (0.55 to 0.80); 0.52 (p=0.003)	8	0.93 (0.66 to 0.99); 3.6 (p=0.08)	6	10 (1.7 to 58)	0.33 (0.23 to 0.49)
• Section thickness >5 mm		0.91 (0.77 to 0.97)	2	Insufficient data	--	--	--
• Excluding studies that used diffusion-weighted imaging		0.76 (0.66 to 0.84); 0.52 (p=0.009)	12	0.95 (0.82 to 0.99); 3.9 (p=0.02)	11	17 (4.2 to 68)	0.25 (0.17 to 0.35)
• Excluding studies restricted to HCC lesions <2 cm		0.92 (0.87 to 0.95); 0.002 (p=0.89)	4	0.88 (0.76 to 0.94); 0.11 (p=0.64)	3	7.6 (3.7 to 16)	0.09 (0.06 to 0.15)
• Excluding studies restricted to hypervascular HCC lesions		0.69 (0.56 to 0.80); 0.43 (p=0.05)	8	0.97 (0.77 to 0.996); 5.5 (p=0.06)	8	20 (2.9 to 141)	0.32 (0.23 to 0.45)
• Histopathological reference standard		0.77 (0.62 to 0.88); 1.0 (p=0.003)	9	0.89 (0.50-0.99); 2.1 (p=0.13)	5	7.1 (1.1 to 46)	0.26 (0.14 to 0.46)
• Histological and clinical/imaging reference standard		0.85 (0.63 to 0.95)	3	0.98 (0.14 to 1.0)	1	46 (0.16 to xxx)	0.15 (0.05 to 0.42)
• Prospective		0.74 (0.57 to 0.86); 0.60 (p=0.005)	6	0.87 (0.56 to 0.97); 2.6 (p=0.01)	5	5.8 (1.5 to 22)	0.30 (0.19 to 0.46)
• United States or Europe		0.74 (0.59 to 0.85); 0.61 (p=0.004)	9	0.96 (0.79 to 0.992); 3.3 (p=0.01)	5	17 (3.5 to 83)	0.27 (0.17 to 0.43)
• Used confidence rating scale		No data	--	No data	--	--	--
• Avoided case-control design		0.79 (0.69 to 0.87); 0.70 (p=0.004)	13	0.95 (0.82 to 0.99); 3.4 (p=0.01)	12	15 (4.4 to 50)	0.22 (0.15 to 0.33)
• Blinded interpretation of imaging		0.85 (0.76 to 0.91); 0.40 (p=0.01)	7	0.94 (0.82 to 0.98); 2.2 (p=0.01)	7	15 (4.7 to 51)	0.16 (0.10-0.25)

HCC = hepatocellular carcinoma; MRI = magnetic resonance imaging

Table 9. Test performance of positron emission tomography for identification of intrahepatic and extrahepatic hepatocellular carcinoma

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
Detection of intrahepatic HCC, FDG PET (KQ 1)	Patient	0.52 (0.39 to 0.66); 0.87 (p=0.01)	15	0.95 (0.82 to 0.99); 0.17 (p=0.40)	5	11 (2.6 to 49)	0.50 (0.37 to 0.68)
• Excluding high risk of bias studies		0.42 (0.26 to 0.60); 0.65 (p=0.01)	7	0.92 (0.76 to 0.98); 0.04 (p=0.73)	3	5.3 (1.5 to 18)	0.63 (0.45 to 0.87)
• PET		0.39 (0.24 to 0.56); 0.54 (p=0.01)	8	0.94 (0.68 to 0.99); 0.11 (p=0.68)	3	6.6 (0.92 to 47)	0.65 (0.48 to 0.88)
• PET/CT		0.65 (0.50-0.78)	7	0.96 (0.74 to 0.99)	2	15 (2.0-111)	0.36 (0.24 to 0.56)
• Prospective		0.46 (0.31 to 0.62); 0.46 (p=0.02)	8	0.94 (0.80-0.98); 0.06 (p=0.62)	4	7.6 (2.0-30)	0.6 (0.41 to 0.79)
• United States or Europe		0.49 (0.32 to 0.66); 0.84 (p=0.01)	10	0.95 (0.82 to 0.99); 0.16 (p=0.41)	5	10 (2.3 to 43)	0.54 (0.37 to 0.77)
• Non-explant histological reference standard		0.46 (0.28 to 0.65); 0.81 (p=0.01)	7	0.91 (0.42 to 0.99); 0.02 (p=0.86)	2	5.2 (0.45 to 60)	0.59 (0.38 to 0.91)
• Mixed histological and imaging/clinical criteria reference standard		0.58 (0.40-0.75)	8	0.97 (0.79 to 0.996)	3	18 (2.4 to 132)	0.43 (0.28 to 0.67)
	HCC lesion	0.56 (0.41 to 0.69); 0.26 (p=0.12)	4	0.91 (0.76 to 0.98)	1	Not calculated	Not calculated
• Excluding high risk of bias studies		0.63 (0.56 to 0.70); <0.0001 (p=1.0)	2	0.91 (0.76 to 0.98)	1	Not calculated	Not calculated
• PET/CT		0.54 (0.37 to 0.70); 0.31 (p=0.18)	3	0.91 (0.76 to 0.98)	1	Not calculated	Not calculated
• Non-explant histological reference standard		0.51 (0.35 to 0.67); 0.25 (p=0.21)	3	No data	No data	--	--
• United States or Europe		0.67 (0.55 to 0.78)	1	0.91 (0.76 to 0.98)	1	Not calculated	Not calculated
Detection of intrahepatic HCC, ^{11}C-acetate PET (KQ 1)	Patient	0.85 (0.67 to 0.94); 0.70 (p=0.13)	4	No data	No data	--	--
	HCC lesion	0.78 (0.61 to 0.89); 0.55 (p=0.15)	4	No data	No data	--	--
• Excluding high risk of bias studies		0.76 (0.69 to 0.82); <0.0001 (p=1.0)	2	No data	No data	--	--
• PET		0.68 (0.46 to 0.84); 0.22 (p=0.49)	2	No data			

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
• PET/CT		0.85 (0.67 to 0.94); 0.42 (p=0.39)	2	No data	No data	--	--
• Non-explant histological reference standard		0.78 (0.61 to 0.89); 0.55 (p=0.15)	4	No data	No data	--	--
• **United States or Europe**		0.78 (0.64 to 0.89)	1				
Detection of intrahepatic HCC, FDG + ^{11}C-acetate PET (KQ 1)	Patient	0.89 (0.80-0.95)	1				
	HCC lesion	0.95 (0.86 to 0.99)	1				
Detection of intrahepatic HCC, ^{18}F-fluorothymidine (KQ 1)	Patient	0.69 (0.41 to 0.89)	1				
Detection of intrahepatic HCC, ^{18}F-fluorochlorine (KQ 1)	Patient	0.91 (0.78 to 0.96); <0.0001 (p=1.0)	2	0.47 (0.23 to 0.72)	1	Not calculated	Not calculated
	HCC lesion	0.84 (0.74 to 0.92)	1	0.62 (0.44 to 0.78)	1	Not calculated	Not calculated
Detection of recurrent intrahepatic HCC, FDG PET (KQ 1)	Patient	0.70 (0.32 to 0.92); 1.5 (p=0.29)	3	0.71 (0.29 to 0.96)	1	Not calculated	Not calculated
	HCC lesion	0.07 (2/27) and 0.73 (22/30)	2	1.0 (1/0)	1	Not calculated	Not calculated
Detection of metastatic HCC, FDG PET (KQ 3)	Patient with metastatic HCC	0.85 (0.71 to 0.93); 0.12 (p=0.13)	6	0.93 (0.89 to 0.95); 1.0 (p=0.17)	5	11 (7.8 to 17)	0.16 (0.08 to 0.33)
• PET		0.98 (0.29 to 0.9998); 6.1 (p=0.30)	2	0.93 (0.86 to 0.97); 0.005 (p=0.85)	2	14 (7.1 to 29)	0.02 (0.0003 to 2.2)
• PET/CT		0.78 (0.72 to 0.83); <0.0001 (p=1.0)	4	0.92 (0.86 to 0.96); <0.0001 (p=1.0)	3	9.8 (5.5 to 17)	0.24 (0.18 to 0.30)
• Excluding high risk of bias studies		0.90 (0.71 to 0.97); 1.1 (p=0.26)	4	0.93 (0.89 to 0.95); <0.0001 (p=1.0)	3	13 (8.3 to 20)	0.11 (0.04 to 0.34)
• Mixed histological and imaging/clinical criteria reference standard		0.78 (0.73 to 0.83); <0.0001 (p=1.0)	5	0.92 (0.87 to 0.96); <0.0001(p=1.0)	4	10 (5.7 to 18)	0.24 (0.18 to 0.30)
• Prospective		0.96 (0.87 to 0.99); <0.0001 (p=1.0)	2	0.93 (0.90-0.96); <0.0001 (p=1.0)	2	14 (9.0-22)	0.05 (0.02 to 0.14)
• United States or Europe		No data	--	No data	--	--	--
• Avoided case-control design		0.88 (0.63 to 0.97); 1.6 (p=0.23)	4	0.92 (0.88 to 0.95); <0.0001 (p=1.0)	3	11 (7.1 to 17)	0.14 (0.04 to 0.47)
• Blinded interpretation of		0.92 (0.80-0.97); 0.26	2	0.93 (0.88 to 0.95)	1	13 (7.8 to	0.08 (0.03 to

	Unit of Analysis	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
imaging		(p=0.23)				20)	0.23)
	Metastatic HCC lesion	0.82 (0.72 to 0.90); 0.17 (p=0.21)	5	No data	--	--	--
• PET		0.81 (0.64 to 0.91); 0.31 (p=0.28)	3	No data	--	--	--
• PET/CT		0.92 (0.77 to 0.97); 0.63 (p=0.31)	3	No data	--	--	--
• Histopathological reference standard		0.85 (0.70-0.93); <0.0001 (p=1.0)	2	No data	--	--	--
• Imaging and clinical criteria		0.90 (0.79 to 0.95)	1	No data	--	--	--
• Mixed histological and imaging/clinical criteria		0.72 (0.58 to 0.82)	2	No data	--	--	--
• Excluding high risk of bias studies		0.86 (0.70-0.95)	1	No data	--	--	--
• Blinded interpretation of imaging		0.87 (0.78 to 0.93); <0.0001 (p=1.0)	3	No data	--	--	--
• United States or Europe		No data[a]	--	No data	--	--	--

CT = computed tomography; FDG = fluorodeoxyglucose; HCC = hepatocellular carcinoma; KQ = Key Question; PET = positron emission tomography

[a] 1 study with 5 patients

Table 10. Pooled direct (within-study) comparisons of test performance of imaging modalities for identification and diagnosis of hepatocellular carcinoma

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Identification of lesions (KQ 1)									
US without contrast (A) vs. CT (B)	Patient	0.68 (0.54 to 0.80)	0.80 (0.68 to 0.88)	-0.12 (-0.20 to -0.03); 0.36 (p=0.15)	6	0.92 (0.84 to 0.96)	0.94 (0.87 to 0.97)	-0.01 (-0.05 to 0.02); 0.62 (p=0.07)	5
• Excluding high risk of bias studies	Patient	0.71 (0.58 to 0.82)	0.82 (0.71 to 0.89)	-0.10 (-0.18 to -0.02); 0.19 (p=0.25)	5	0.91 (0.81 to 0.96)	0.94 (0.86 to 0.98)	-0.03 (-0.07 to 0.01); 0.73 (p=0.11)	4
US without contrast (A) vs. CT (B)	Lesion	0.55 (0.43 to 0.66)	0.66 (0.54 to 0.76)	-0.11 (-0.18 to -0.04); 0.11 (p=0.28)	3	0.83 (0.65 to 0.93)	0.93 (0.83 to 0.98)	-0.10 (-0.20 to -0.008); 0.44 (p=0.29)	2
• HCC lesions <2 cm	Lesion	0.46 (0.30-0.63)	0.54 (0.37 to 0.70)	-0.07 (-0.17 to 0.02); 0.31 (p=0.27)	3	0.72 (0.61 to 0.80)	0.80 (0.71 to 0.86)	-0.08 (-0.20 to 0.04); 0.002 (p=0.85)	2
US with contrast (A) vs. CT (B)	Lesion	0.58 (0.37 to 0.77)	0.74 (0.54 to 0.87)	-0.16 (-0.32 to -0.01); 0.50 (p=0.15)	3	No data	No data	--	--
• HCC lesions <2 cm	Lesion	0.30 (0.17 to 0.43)	0.44 (0.30-0.58)	-0.14 (-0.32 to 0.05)	1	No data	No data	--	--
US without contrast (A) vs. MRI (B)	Patient	0.61 (0.48 to 0.74)	0.81 (0.69 to 0.89)	-0.19 (-0.30 to -0.08); 0.01 (p=0.79)	3	0.94 (0.87 to 0.97)	0.82 (0.66 to 0.91)	0.13 (0.03 to 0.22); 0.01 (p=0.40)	3
US without contrast (A) vs. MRI (B)	Lesion	0.57 (0.42 to 0.71)	0.79 (0.67 to 0.88)	-0.22 (-0.31 to -0.14); 0.22 (p=0.28)	3	0.75 (0.66 to 0.82)	0.78 (0.70-0.85)	-0.03 (-0.13 to 0.06); 0.001 (p=0.89)	2
• HCC lesion <2 cm	Lesion	0.40 (0.18 to 0.67)	0.65 (0.38 to 0.85)	-0.26 (-0.36 to -0.15); 0.60	2	0.71 (0.60-0.80)	0.84 (0.76 to 0.89)	-0.13 (-0.25 to -0.01); 0.006	2
US with contrast (A) vs. MRI (B)	Lesion	0.54 (0.25 to 0.80)	0.70 (0.40-0.89)	-0.16 (-0.30 to -0.02); 0.71 (p=0.31)	2	No data	No data	--	--
• HCC lesions <2 cm	Lesion	0.30 (0.17 to 0.43)	0.42 (0.28 to 0.56)	-0.12 (-0.31 to 0.07)	1	No data	No data	--	--

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
• Well-differentiated HCC Lesion	Lesion	0.43 (0.14 to 0.77)	0.36 (0.11 to 0.72)	0.07 (-0.19 to 0.33); 0.87 (p=0.34)	2	No data	No data	--	--
MRI (A) vs. CT (B)	Patient	0.88 (0.53 to 0.98)	0.82 (0.41 to 0.97)	0.06 (-0.05 to 0.17); 3.0 (p=0.21)	4	0.84 (0.70-0.92)	0.91 (0.82 to 0.96)	-0.08 (-0.16 to 0.00); 0.40 (p=0.21)	4
• Excluding high risk of bias studies	Patient	0.82 (0.75 to 0.88)	0.75 (0.68 to 0.81)	0.07 (-0.02 to 0.17); <0.0001	2	0.80 (0.57 to 0.92)	0.91 (0.77 to 0.97)	-0.11 (-0.23 to 0.01); 0.44	2
MRI (A) vs. CT (B)	Lesion	0.81 (0.77 to 0.84)	0.72 (0.67 to 0.77)	0.09 (0.06 to 0.12); 0.37 (p<0.0001)	28	0.85 (0.76 to 0.92)	0.90 (0.82 to 0.95)	-0.05 (-0.10 to 0.01); 0.43 (p=0.01)	6
• Excluding high risk of bias studies	Lesion	0.80 (0.73 to 0.85)	0.73 (0.66 to 0.79)	0.07 (0.04 to 0.10); 0.50 (p<0.0001)	19	0.87 (0.78 to 0.93)	0.93 (0.86 to 0.96)	-0.05 (-0.10 to 0.00); 0.37 (p=0.03)	5
• Non-hepatic specific contrast	Lesion	0.81 (0.74 to 0.87)	0.74 (0.65 to 0.81)	0.08 (0.04 to 0.11); 0.39 (p=0.01)	11	0.62 (0.51 to 0.72)	0.86 (0.77 to 0.93)	-0.24 (-0.37 to -0.11); <0.0001 (p=1.0)	2
• Hepatic specific contrast	Lesion	0.80 (0.74 to 0.85)	0.70 (0.62 to 0.77)	0.10 (0.06 to 0.14); 0.41 (p=0.0003)	15	0.93 (0.88 to 0.96)	0.91 (0.85 to 0.94)	0.02 (-0.03 to 0.07); 0.01 (p=0.78)	4
• HCC lesions <2 cm	Lesion	0.59 (0.43 to 0.73)	0.46 (0.32 to 0.62	0.12 (0.03 to 0.22); 0.25 (p=0.27)	3	0.84 (0.73 to 0.91)	0.80 (0.67 to 0.89)	0.04 (-0.06 to 0.14); 0.12 (p=0.41)	2
Evaluation of previously identified lesion (KQ 2)									
US without contrast (A) vs. CT (B)	Patient	0.78 (0.70-0.85)	0.89 (0.84 to 0.95)	-0.12 (-0.21 to -0.02)	1	No data	No data	--	--
US with contrast (A) vs. CT (B)	Patient	0.91 (0.85 to 0.95)	0.87 (0.79 to 0.92)	0.04 (-0.01 to 0.10); 0.17 (p=0.18)	4	0.93 (0.88 to 0.97)	0.94 (0.88 to 0.97)	-0.01 (-0.06 to 0.05); 0.07 (p=0.36)	2
• Excluding high risk of bias studies	Patient	0.90 (0.76 to 0.96)	0.84 (0.68 to 0.93)	0.06 (-0.03 to 0.15); 0.32 (p=xxxx)	2	0.94 (0.87 to 0.97)	0.94 (0.88 to 0.97)	-0.01 (-0.06 to 0.05); 0.08 (p=xxxx)	2
US with contrast (A) vs. CT (B)	Lesion	0.94 (0.89 to 0.97)	0.91 (0.85 to 0.94)	0.03 (-0.03 to 0.09); <0.0001 (p=1.0)	3	No data	No data	--	--

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
• HCC lesion <2 cm	Lesion	0.78 (0.61 to 0.89)	0.71 (0.52 to 0.85)	0.07 (-0.01 to 0.15); 1.1 (p-0.02)	7	0.87 (0.62 to 0.97)	0.94 (0.77 to 0.98)	-0.06 (-0.15 to 0.03); 2.4 (p=0.09)	4
• Well-differentiated HCC Lesion	Lesion	0.55 (0.25 to 0.82)	0.55 (0.25 to 0.82)	0.00 (-0.30 to 0.30); 0.48 (p=0.40)	2	No data	No data	--	--
US with contrast (A) vs. MRI (B)	Patient	0.79 (0.65 to 0.94)	0.83 (0.69 to 0.97)	-0.03 (-0.24 to 0.17)	1	0.79 (0.68 to 0.90)	0.75 (0.64 to 0.87)	0.04 (-0.12 to 0.20)	1
• HCC lesion <2 cm	Patient	0.52 (0.39 to 0.64)	0.62 (0.49 to 0.74)	-0.10 (-0.27 to 0.08)	1	0.93 (0.84 to 1.0)	0.97 (0.90-1.0)	-0.03 (-0.15 to 0.08)	1
US with contrast (A) vs. MRI (B)	Lesion	0.79 (0.65 to 0.94)	0.83 (0.69 to 0.97)	-0.03 (-0.24 to 0.17)	1	0.79 (0.68 to 0.90)	0.75 (0.64 to 0.87)	0.04 (-0.12 to 0.20)	1
• HCC lesion <2 cm	Lesion	0.53 (0.28 to 0.76)	0.68 (0.43 to 0.86)	-0.16 (-0.30 to -0.02); 0.72 (p=0.25)	3	0.95 (0.85 to 0.98)	0.98 (0.91 to 0.99)	-0.03 (-0.08 to 0.02); 0.38 (p=0.43)	3
MRI (A) vs. CT (B)	Patient	0.81 (0.70-0.92)	0.74 (0.62 to 0.87)	0.06 (-0.10 to 0.23)	1	0.85 (0.72 to 0.99)	0.81 (0.66 to 0.96)	0.04 (-0.16 to 0.24)	1
MRI (A) vs. CT (B)	Lesion	0.84 (0.76 to 0.92)	0.62 (0.52 to 0.72)	0.22 (0.09 to 0.35)	1	0.36 (0.20-0.52)	0.72 (0.58 to 0.87)	-0.36 (-0.58 to -0.15)	1
Identification of metastatic HCC (KQ 4)									
PET/CT (A) vs. CT (B)	Patient (2), lesion (1)	0.82 (0.61 to 0.93)	0.85 (0.66 to 0.95)	-0.03 (-0.12 to 0.060); 0.75 (p=0.17)	3	Insufficient data	Insufficient data	--	--

CT = computed tomography; HCC = hepatocellular carcinoma; MRI =magnetic resonance imaging; PET = positron emission tomography; US = ultrasound

Table 11. Comparisons of test performance for hepatocellular carcinoma of single compared with multiple modality imaging[a]

Study, Year	Unit of Analysis	Single Imaging Modalities	Sensitivity	Specificity	Multiple Imaging Modalities	Criteria for Positive Results with Multiple Imaging Modalities	Sensitivity	Specificity
Identification of HCC								
Alaboudy, 2011[48]	Lesion	A: US B: CT C: MRI	A: 0.72 B: 0.74 C: 0.86	Not reported	A: US + MRI B: US + CT C: CT + MRI	Unclear	A: 0.90 B: 0.82 C: 0.88	Not reported
Iavarone, 2010[64]	Lesion (1 to 2 cm)	A: MRI B: CT C: US	A: 0.42 B: 0.45 C: 0.32	Not reported	A: US + MR B: US + CT C: MRI + CT D: Any dual combination of MRI, CT, and US	Concordant positive findings on 2 imaging modalities	A: 0.16 B: 0.19 C: 0.29 D: 0.40	Not reported
Diagnosis of HCC								
Dai, 2008[52]	Lesion	A: US B: CT	A: 0.91 B: 0.80	A: 0.87 B: 0.98	CT + US	Unclear	0.80	0.87
Forner, 2008[55]	Lesion (<2 cm)	MRI	0.62	0.97	MRI + US	1: Definite positive findings on 2 imaging modalities 2: “At least suspicious” on 2 imaging modalities	1: 0.33 2: 0.67	1: 1.0 2: 1.0
Golfieri, 2009[121]	Lesion (<3 cm)	CT	0.62	0.72	MRI + CT	Unclear	0.89	0.22
Khalili, 2011[69]	Lesion (1 to 2 cm)	CT	0.53	0.99	A: US + MRI B: CT + MRI C: CT + US D: MRI then US E: MRI then CT F: CT then US	A-C: Concordant positive results on 2 imaging modalities D-F: Positive findings on initial imaging modality or positive findings on second imaging modality for indeterminate findings on first scan	A: 0.35 B: 0.41 C: 0.29 D: 0.79 E: 0.74 F: 0.76	A: 1.0 B: 1.0 C: 0.99 D: 0.91 E: 0.99 F: 0.91
Quaia, 2009[87]	Lesion (<3 cm)	A: CT B: US	A: 0.72 B: 0.88	A: 0.71 B: 0.66	CT + US	Positive findings from at least one imaging technique	0.97	0.70

Study, Year	Unit of Analysis	Single Imaging Modalities	Sensitivity	Specificity	Multiple Imaging Modalities	Criteria for Positive Results with Multiple Imaging Modalities	Sensitivity	Specificity
Sangiovanni, 2010[42]	Lesion (1 to 2 cm)	A: US B: CT C: MRI	A: 0.26 B: 0.44 C: 0.44	A: 1.0 B: 1.0 C: 1.0	US, CT, and MRI	1: Concordant positive findings on two imaging techniques 2: Positive findings from at least one imaging technique	1: 0.35 2: 0.65	Not reported
Serste, 2012[43]	Patient	CT	0.74	0.81	CT + MRI	1: Concordant positive findings on two imaging techniques 2: Positive findings from at least one imaging technique	1: 0.57 2: 0.98	1: 0.85 2: 0.81

CT = computed tomography; HCC = hepatocellular carcinoma; MRI = magnetic resonance imaging; US = ultrasound

[a] Ultrasound contrast-enhanced in all studies except Forner 2008 and Iavarone 2010

Table 12. Pooled direct (within-study) comparisons of ultrasound for identification and diagnosis of hepatocellular carcinoma

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Identification of HCC lesions									
Contrast (A) vs. no contrast (B)	Lesion (1), liver segment (1)	0.79 (0.72 to 0.76))	0.81 (0.76 to 0.86)	-0.04 (-.11 to 0.04); -0.04 (p=1.0)	2	0.98 (0.96 to 0.997)	0.92 (0.89 to 0.95)	0.06 (0.02 to 0.10)	1
Doppler (A) vs. no Doppler B)	Lesion	0.67 (0.52 to 0.81)	0.60 (0.45 to 0.74)	0.07 (-0.13 to 0.28)	1	No data	No data	--	--
Diagnosis of HCC lesions (previously identified liver lesions)									
Contrast (A) vs. no contrast (B)	Lesion	0.89 (0.83 to 0.93)	0.39 (0.32 to 0.47)	0.50 (0.41 to 0.58); <0.0001 (p=1.0)	2	1.0 (1.0-1.0)	0.94 (0.85 to 1.0)	0.06 (-0.02 to 0.15)	1
Doppler (A) vs. no Doppler (B)	Lesion	0.69 (0.29 to 0.93)	0.68 (0.28 to 0.92)	-0.01 (-0.15 to 0.13); 1.2 (p=0.34)	2	No data	No data	--	--
Doppler (A) vs. no Doppler (B) (also contrast vs. no contrast)	Patient	0.93 (0.88 to 0.97)	0.78 (0.70-0.85)	0.15 (0.06 to 0.23)	1	No data	No data	--	--
Moderately or poorly-differentiated (A) vs. well-differentiated HCC lesion (B), with contrast	Lesion	0.83 (0.55 to 0.95)	0.43 (0.15 to 0.76)	0.40 (0.17 to 0.64); 1.2 (p=0.17)	3	No data	No data	--	--
HCC lesion >=20 mm (A) vs. <20 mm (B)	Patient (1), lesion (12), liver segment (1)	0.88 (0.78 to 0.94)	0.49 (0.31 to 0.67)	0.39 (0.27 to 0.51); 1.9 (p=0.0004)	14	0.89 (0.48 to 0.99)	0.91 (0.53 to 0.99)	-0.01 (-0.09 to 0.06); 2.8 (p=0.21)	3
• No contrast	Patient (1), lesion (7), liver segment (1)	0.82 (0.68 to 0.91)	0.34 (0.19 to 0.53)	0.48 (0.39 to 0.57); 1.3 (p=0.005)	9	0.80 (0.61 to 0.91)	0.81 (0.62 to 0.92)	-0.01 (-0.13 to 0.11); 0.51 (p=0.08)	2
• With contrast	Lesion	0.94 (0.83 to 0.98)	0.77 (0.53 to 0.91)	0.17 (0.03 to 0.32); 1.3 (p=0.05)	5	1.0 (26/26)	1.0 (2/2)	Not calculated	1

HCC = hepatocellular carcinoma

Table 13. Direct comparisons of diagnostic accuracy according to lesion size <10, 10-20, and >20 mm

	Sensitivity (95% CI); τ^2 (p value)	Number of Studies	Specificity (95% CI); τ^2 (p value)	Number of Studies	LR+	LR-
Ultrasound without contrast						
• <10 mm	0.09 (0.02 to 0.29); 1.5 (p=0.08)	4	0.93 (0.79 to 1.0)	1	1.3	0.98
• 10-20 mm	0.50 (0.23 to 0.78)	4	0.60 (0.46 to 0.74)	1	1.2	0.83
• >20 mm	0.88 (0.66 to 0.96)	4	0.53 (0.35 to 0.71)	1	1.9	0.23
• Difference >20 mm vs. 10-20 mm	0.37 (0.18 to 0.57)	4	-0.33 (-0.53 to -0.14)	1	--	--
• Difference 10-20 mm vs. <10 mm	0.41 (0.19 to 0.63)	4	-0.06 (-0.29 to 0.16)	1	--	--
Ultrasound with contrast						
• 10-20	0.64 (0.33 to 0.87); 1.2 (p=0.15)	3	1.0 (26/26)	1	--	--
• >20 mm	0.91 (0.71 to 0.98)	3	1.0 (2/2)	1	--	--
• Difference >20 mm vs. 10-20 mm	0.26 (0.04 to 0.48)	3	0.0	0	--	--
Computed Tomography (CT)						
• <10 mm	0.32 (0.24 to 0.40); 0.51 (p<0.0001)	20	0.69 (0.52 to 0.82); <0.0001 (p=0.9998)	2	1.0 (0.58 to 1.8)	0.99 (0.77 to 1.3)
• 10-20 mm	0.73 (0.66 to 0.80)	22	0.86 (0.74 to 0.93)	2	5.2 (2.7 to 10)	0.31 (0.23 to 0.41)
• >20 mm	0.95 (0.92 to 0.97)	19	0.90 (0.73 to 0.97)	1	9.5 (3.2 to 28)	0.90 (0.73 to 0.97)
• Difference >20 mm vs. 10-20 mm	0.42 (0.35 to 0.48)	21	0.17 (-0.02 to 0.36)	1	--	--
• Difference 10-20 mm vs. <10 mm	0.21 (0.15 to 0.27)	20	0.04 (-0.11 to 0.19)	2	--	--
Magnetic Resonance Imaging (MRI)						
• <10 mm	0.43 (0.32 to 0.54); 0.72 (p<0.0001)	19	0.69 (0.23 to 0.94); 2.9 (p=0.01)	2	1.4 (0.35 to 5.5)	0.83 (0.45 to 1.5)
• 10-20 mm	0.77 (0.67 to 0.84)	18	0.84 (0.42 to 0.97)	2	4.8 (0.i93 to 24)	0.28 (0.18 to 0.43)
• >20 mm	0.97 (0.94 to 0.98)	14	0.93 (0.57 to 0.99)	2	13 (1.7 to 101)	0.03 (0.02 to 0.07)
• Difference >20 mm vs. 10-20 mm	0.20 (0.13 to 0.28)	13	0.09 (-0.09 to 0.27)	2	--	--
• Difference 10-20 mm vs. <10 mm	0.34 (0.27 to 0.41)	17	0.15 (-0.11 to 0.40)	2	--	--

Table 14. Computed tomography direct comparisons

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Section thickness 7. 5 (A) vs. 5 mm (B)	Lesion	0.64 (0.58 to 0.70)	0.72 (0.64 to 0.78)	-0.07 (-0.17 to 0.02); <0.0001 (p=1.0)	2	No data	No data	--	--
Section thickness 7.5 (A) vs. 5 mm (B), restricted to lesions <2 cm	Lesion	0.39 (0.27 to 0.52)	0.41 (0.22 to 0.59)	-0.01 (-0.24 to 0.21)	1	No data	No data	--	--
Spectral CT (A) vs. standard CT (B)	Lesion	0.97 (0.89 to 0.99)	0.91 (0.80-0.97)	0.05 (-0.02 to 0.12); 0.33 (p=0.39)	3	0.98 (0.80-0.998)	0.92 (0.64 to 0.99)	0.06 (-0.06 to 0.18); 0.99 (p=0.59)	2
Moderately or poorly (A) vs. well differentiated (B) HCC lesion	Lesion	0.82 (0.66 to 0.91)	0.50 (0.29 to 0.70)	0.32 (0.19 to 0.45); 0.77 (p=0.05)	5	No data	No data	--	--
HCC lesion >=20 mm (A) vs. <20 mm (B)	Lesion (30); liver segment (3)	0.94 (0.91 to 0.95)	0.62 (0.56 to 0.68)	0.31 (0.26 to 0.36); 0.50 (p<0.0001)	33	0.92 (0.85 to 0.96)	0.80 (0.71 to 0.86)	0.12 (0.03 to 0.21); 0.50 (p=1.0)	2
HCC lesion >=20 mm (A) vs. <20 mm (B), restricted to studies meeting minimum technical criteria*	Lesion	0.94 (0.89 to 0.97)	0.60 (0.49 to 0.70)	0.35 (0.25 to 0.44); 0.26 (p=0.04)	7	0.90 (0.73 to 0.97)	0.85 (0.74 to 0.92)	0.05 (-0.09 to 0.19)	1
Cirrhosis (A) vs. no cirrhosis (B)	Lesion	0.85 (0.77 to 0.91)	0.81 (0.74 to 0.87)	0.04 (-0.05 to 0.14); <0.0001 (p=1.0)	2	No data	No data	--	--

HCC = hepatocellular carcinoma

Table 15. MRI direct comparisons

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Gadoxetic acid or gadobenate (A) vs. gadopentetate or gadodiadmide (B)	Lesion	0.82 (0.71 to 0.90)	0.75 (0.61 to 0.85)	0.07 (0.01 to 0.14); 0.43 (p=0.06)	5	0.92 (0.81 to 0.96)	0.92 (0.82 to 0.97)	-0.004 (-0.10-0.09); 0.002 (p=0.93)	2
Gadoxetic acid or gadobenate (A) vs. gadopentetate or gadodiadmide (B), for HCC lesions <2 cm	Lesion	0.77 (0.68 to 0.84)	0.62 (0.52 to 0.71)	0.15 (0.08 to 0.22); 0.20 (p=0.05)	7	0.93 (0.82 to 0.98	0.91 (0.79 to 0.97)	0.02 (-0.05 to 0.09); 0.14 (p=0.51)	2
Diffusion-weighted imaging (A) vs. no diffusion-weighted imaging (B)	Lesion (7), patient (1)	0.81 (0.74 to 0.86)	0.81 (0.75 to 0.86)	-0.01 (-0.05 to 0.03); 0.14 (p=0.05)	8	0.92 (0.83 to 0.97)	0.81 (0.65 to 0.91)	0.11 (0.02 to 0.20); 0.73 (p=0.13)	5
Diffusion-weighted imaging (A) vs. no diffusion-weighted imaging (B) for HCC lesion <2 cm	Lesion	0.78 (0.62 to 0.88)	0.67 (0.50-0.81)	0.10 (0.02 to 0.18); 0.75 (p=0.03)	5	0.97 (0.31 to 0.9995)	0.91 (0.15 to 0.999)	0.06 (-0.16 to 0.28); 4.4 (p=0.37)	2
Moderately or poorly (A) vs. well differentiated (B) HCC lesion	Lesion	0.54 (0.26 to 0.79)	0.38 (0.17 to 0.64)	0.16 (-0.11 to 0.43); 0.39 (p=0.40)	2	No data	No data	--	--

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
HCC lesion >=20 mm (A) vs. <20 mm (B)	Lesion (23); liver segment (1); patient (1)	0.96 (0.93 to 0.97)	0.65 (0.57 to 0.73)	0.30 (0.23 to 0.37); 0.72 (p<0.0001)	25	0.96 (0.83 to 0.99)	0.88 (0.61 to 0.97)	0.08 (-0.03 to 0.20)	5
HCC lesion >=20 mm (A) vs. <20 mm (B), gadopentetate or gadodiamide contrast	Lesion (13); liver segment (1); patient (1)	0.94 (0.90-0.96)	0.54 (0.43 to 0.65)	0.40 (030-0.49); 0.63 (p=0.0005)	15	0.98 (0.89 to 0.998)	0.94 (0.73 to 0.99)	0.04 (-0.03 to 0.11); 2.5 (p=0.07)	3
HCC lesion >=20 mm (A) vs. <20 mm (B), gadoxetic acid or gadobenate contrast	Lesion (9)	0.97 (0.94 to 0.99)	0.77 (0.69 to 0.84)	0.20 (0.13 to 0.27); 0.28 (p=0.02)	9	0.93 (0.77 to 0.98)	0.90 (0.80-0.96)	0.03 (-0.08 to 0.15)	1
HCC lesion >=20 mm (A) vs. <20 mm (B), restricted to studies meeting minimum technical criteria*	Lesion	0.94 (0.89 to 0.97)	0.60 (0.49 to 0.70)	0.35 (0.25 to 0.44); 0.26 (p=0.04)	7	0.90 (0.73 to 0.97)	0.85 (0.74 to 0.92)	0.05 (-0.09 to 0.19)	1

HCC = hepatocellular carcinoma

Table 16. Positron emission tomography direct comparisons

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Identification of intrahepatic HCC (KQ 1)									
FDG (A) vs. ^{11}C-acetate (B) PET	Patient	0.58 (0.44 to 0.70)	0.81 (0.70-0.89)	-0.23 (-0.34 to -0.13) ; 0.13 (p=0.23)	3	No data	No data	--	--
	Lesion	0.52 (0.45 to 0.59)	0.79 (0.72 to 0.84)	-0.27 (-0.36 to -0.17)	3	No data	No data	--	--
FDG (A) vs. FDG + ^{11}C-acetate (B) PET	Patient	0.63 (0.52 to 0.74)	0.89 (-0.83 to 0.96)	-0.26 (-0.39 to -0.13)	1	No data	No data	--	--
	Lesion	0.33 (0.21 to 0.45)	0.95 (0.89 to 1.0)	-0.62 (-0.75 to -0.49)	1	No data	No data	--	--
FDG (A) vs. ^{18}F-fluorocholine	Patient	0.65 (0.50-0.78)	0.91 (0.78 to 0.96)	-0.26 (-0.42 to -0.09)	2	0.94 (0.83 to 1.0)	0.47 (0.23 to 0.71)	0.47 (0.23 to 0.71)	1
	Lesion	0.67 (0.56 to 0.78)	0.84 (0.76 to 0.93)	-0.17 (-0.31 to -0.03)	1	0.91 (0.82 to 1.0)	0.62 (0.45 to 0.78)	0.29 (0.10-0.48)	1
Well-differentiated vs. moderately or poorly-differentiated HCC lesion, FDG PET	Patient (2), lesion (3)	0.49 (0.41 to 0.58)	0.78 (0.70-0.85)	-0.29 (-0.58 to -0.41); <0.0001 (p=1.0)	5				
Identification of metastatic HCC (KQ 3)									
FDG (A) vs. ^{11}C-acetate (B) PET	Patient	0.79 (0.71 to 0.87)	0.64 (0.54 to 0.73)	0.15 (0.03 to 0.28)	1	0.91 (0.79 to 1.0)	0.95 (0.87 to 1.0)	-0.05 (-0.19 to 0.10)	1
FDG (A) vs. FDG + ^{11}C-acetate (B) PET	Patient	0.79 (0.71 to 0.87)	0.98 (0.95 to 1.0)	-0.19 (-0.28 to -0.11)	1	0.91 (0.79 to 1.0)	0.86 (0.72 to 1.0)	0.05 (-0.14 to 0.23)	1
PET vs. PET/CT	Metastatic HCC lesion	0.90 (0.82 to 0.97)	0.98 (0.95 to 1.0)	-0.09 (-0.17 to 0.0)	1	No data	No data	--	--

CT = computed tomography; FDG = fludeoxyglucose; HCC = hepatocellular carcinoma; PET = positron emission tomography

Table 17. Test performance of fludeoxyglucose positron emission tomography for hepatocellular carcinoma, stratified by lesion size

Study, Year	HCC Location	Unit of Analysis	Sensitivity: FDG	Sensitivity: ^{11}C-acetate
Cheung, 2011[272]	Intrahepatic	Patient	≤5 cm: 0.24 (7/29) >5 cm: 0.62 (18/29)	≤5 cm: 0.97 (28/29) >5 cm: 0.97 (28/29)
Kim YK, 2010 (3)[279]	Intrahepatic	Lesion	<1 cm: 0.0 (0/21) ≥1 cm: 0.33 (2/6)	--
Park JW, 2008[283]	Intrahepatic	Lesion	≥1 to 2 cm: 0.27 (6/22) ≥2 to 5 cm: 0.48 (22/46) ≥5 cm: 0.93 (39/42)	≥1 to 2 cm: 0.32 (7/22) ≥2 to 5 cm: 0.78 (36/46) ≥5 cm: 0.95 (40/42)
Trojan, 1999[99]	Intrahepatic	Patient	<5 cm: 0.12 (1/8) ≥5 cm: 1.0 (6/6)	--
Wolfort, 2010[290]	Intrahepatic	Lesion	≤5 cm: 0.25 (2/8) >5 cm: 1.0 (5/5)	--
Kim YK, 2010 (3)[279]	Extrahepatic	Lesion	<1 cm: 0.0 (0/2) ≤1 cm: 0.93 (13/14)	--
Lee JE, 2012[156]	Lung	Patient	<1 cm: 0.20 (2/10) ≥1 cm: 0.92 (12/13)	--
Park JW, 2008[283]	Extrahepatic	Lesion	<=1 to 2 cm: 0.80 (16/20) ≥2 cm: 0.93 (14/15)	≤1 to 2 cm: 0.65 (13/20) ≥2 cm: 0.93 (14/15)
Sugiyama, 2004[285]	Extrahepatic	Lesion	<1 cm: 0.12 (1/8) ≥1 cm: 0.83 (24/29)	--

FDG = fludeoxyglucose; HCC = hepatocellular carcinoma

Table 18. Test performance of fludeoxyglucose positron emission tomography for hepatocellular carcinoma, stratified by lesion size

Study, Year	HCC Location	Unit of Analysis	Sensitivity: FDG	Sensitivity: ^{11}C-acetate
Cheung, 2011[272]	Intrahepatic	Patient	≤5 cm: 0.24 (7/29) >5 cm: 0.62 (18/29)	≤5 cm: 0.97 (28/29) >5 cm: 0.97 (28/29)
Kim YK, 2010 (3)[279]	Intrahepatic	Lesion	<1 cm: 0.0 (0/21) ≥1 cm: 0.33 (2/6)	--
Park JW, 2008[283]	Intrahepatic	Lesion	≥1 to 2 cm: 0.27 (6/22) ≥2 to 5 cm: 0.48 (22/46) ≥5 cm: 0.93 (39/42)	≥1 to 2 cm: 0.32 (7/22) ≥2 to 5 cm: 0.78 (36/46) ≥5 cm: 0.95 (40/42)
Trojan, 1999[99]	Intrahepatic	Patient	<5 cm: 0.12 (1/8) ≥5 cm: 1.0 (6/6)	--
Wolfort, 2010[290]	Intrahepatic	Lesion	≤5 cm: 0.25 (2/8) >5 cm: 1.0 (5/5)	--
Kim YK, 2010 (3)[279]	Extrahepatic	Lesion	<1 cm: 0.0 (0/2) ≤1 cm: 0.93 (13/14)	--
Lee JE, 2012[156]	Lung	Patient	<1 cm: 0.20 (2/10) ≥1 cm: 0.92 (12/13)	--
Park JW, 2008[283]	Extrahepatic	Lesion	<=1 to 2 cm: 0.80 (16/20) ≥2 cm: 0.93 (14/15)	≤1 to 2 cm: 0.65 (13/20) ≥2 cm: 0.93 (14/15)
Sugiyama, 2004[285]	Extrahepatic	Lesion	<1 cm: 0.12 (1/8) ≥1 cm: 0.83 (24/29)	--

FDG = fludeoxyglucose; HCC = hepatocellular carcinoma

Table 19. Studies on accuracy of imaging for differentiating hepatocellular carcinoma from other lesions

Study, Year	Imaging Modality	HCC Lesion	Lesion for Differentiation	Unit of Analysis	Reference Standard	Sensitivity	Specificity	Diagnostic Criteria
Pei, 2012[86]	US with contrast	Hypervascular HCC	Focal nodular hyperplasia	Patient	Histological	0.94 (62/66)	0.68 (23/34)	Based on quantitative analysis of contrast-enhanced US findings
Kim SE, 2011[141]	CT	HCC	Non-HCC lesion (including cholangiocarcioma, metastasis, and FNH)	Lesion	Histological	0.85 (140/164)	0.90 (38/42)	Based on arterial enhancement and venous washout
Lv, 2011[166]	CT	Hypervascular HCC lesion <3 cm	Hemangioma	Lesion	Mixed histological and clinical/imaging	0.91 (32/35)	0.87 (26/30)	Based on enhancement pattern on standard CT images
Sun, 2010[191]	CT	Hypervascular HCC lesion <2 cm	Hypervascular pseudolesion	Patient	Mixed histological and clinical/imaging	0.54 (18/33)	0.96 (26/27)	Confidence level score of 4-5 on 1 to 5 scale, based on enhancement pattern
Yu, 2013[196]	CT	HCC	Focal nodular hyperplasia	Lesion	Histological	0.95 (40/42)	1.0 (16/16)	Criteria for diagnosis not defined
Ito, 2004[218]	MRI	Hypervascular HCC lesion <3 cm	Hypervascular pseudolesion	Lesion	Mixed histological and clinical/imaging	0.52 (21/40)	1.0 (30/30)	Based on rapid central washout after early enhancement and peritumoral coronal enhancement
Jeong, 1999[219]	MRI	Hypervascular HCC lesion	Hemangioma	Lesion	Mixed histological and clinical/imaging	0.94 (31/33)	0.82 (15/18)	Based on contrast to noise ratio of 7.00 on imaging 60 s after administration of contrast
Kamura, 2002[222]	MRI	Hypervascular HCC lesion <2 cm	Hypervascular pseudolesion	Lesion	Mixed histological and clinical/imaging	0.47 (9/19)	0.93 (13/14)	Based on hyperintensity on T2 to weighted images
Lee MH, 2011[243]	MRI	Well-differentiated HCC lesion	Benign nodule (regenerative nodule or dysplastic nodule)	Lesion	Histological	0.85 (39/46)	0.42 (10/24)	Based on hypointensity on hepatobiliary phase imaging
Motosugi, 2010[247]	MRI	Hypervascular HCC lesion (mean 16 mm)	Hypervascular pseudolesion (mean 11 mm)	Lesion	Mixed histological and clinical/imaging	0.91 (112/123)	0.91 (29/32)	Based on hepatocyte-phase signal intensity ratio of 0.84 on gadoxetic-enhanced images
Sun, 2010[191]	MRI	Hypervascular HCC lesion <2 cm	Hypervascular pseudolesion	Patient	Mixed histological and clinical/imaging	0.92 (30.5/33)	0.94 (25/27)	Confidence level score 4-5 on 1 to 5 scale, based on enhancement pattern

Study, Year	Imaging Modality	HCC Lesion	Lesion for Differentiation	Unit of Analysis	Reference Standard	Sensitivity	Specificity	Diagnostic Criteria
Vandecaveye, 2009[264]	MRI	HCC lesion[a]	Benign lesions (regenerative nodules, low-grade dysplastic nodules, stable lesions, or other benign lesions)	Lesion	Mixed histological and clinical/imaging	0.81 (50/62)	0.65 (34/52)	Based on T2- and T1-signal intensity ratio and enhancement pattern
Xu, 2010[266]	MRI	HCC lesion	Dysplastic nodule	Lesion	Histological	0.82 (33/40)	0.58 (11/18)	Confidence level score 4-5 on 1 to 5 scale, based on enhancement pattern
Yu, 2002[269]	MRI	HCC lesion <4 cm	Cavernous hemangioma	Lesion	Mixed histological and clinical/imaging	0.88 (137/155)	0.15 (31/207)	Based on absence of transient peritumoral enhancement

CT = computed tomography; HCC = hepatocellular carcinoma; MRI = magnetic resonance imaging; US = ultrasound

[a] Including two cholangiocarcinomas and two high-grade dysplastic nodules

Table 20. Pooled direct (within-study) comparisons of test performance of imaging modalities for identification and diagnosis of hepatocellular carcinoma

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
Identification of lesions (KQ 1)									
US without contrast (A) vs. CT (B)	Patient	0.68 (0.54 to 0.80)	0.80 (0.68 to 0.88)	-0.12 (-0.20 to -0.03); 0.36 (p=0.15)	6	0.92 (0.84 to 0.96)	0.94 (0.87 to 0.97)	-0.01 (-0.05 to 0.02); 0.62 (p=0.07)	5
• Excluding high risk of bias studies	Patient	0.71 (0.58 to 0.82)	0.82 (0.71 to 0.89)	-0.10 (-0.18 to -0.02); 0.19 (p=0.25)	5	0.91 (0.81 to 0.96)	0.94 (0.86 to 0.98)	-0.03 (-0.07 to 0.01); 0.73 (p=0.11)	4
US without contrast (A) vs. CT (B)	Lesion	0.55 (0.43 to 0.66)	0.66 (0.54 to 0.76)	-0.11 (-0.18 to -0.04); 0.11 (p=0.28)	3	0.83 (0.65 to 0.93)	0.93 (0.83 to 0.98)	-0.10 (-0.20 to -0.008); 0.44 (p=0.29)	2
• HCC lesions <2 cm	Lesion	0.46 (0.30-0.63)	0.54 (0.37 to 0.70)	-0.07 (-0.17 to 0.02); 0.31 (p=0.27)	3	0.72 (0.61 to 0.80)	0.80 (0.71 to 0.86)	-0.08 (-0.20 to 0.04); 0.002 (p=0.85)	2
US with contrast (A) vs. CT (B)	Lesion	0.58 (0.37 to 0.77)	0.74 (0.54 to 0.87)	-0.16 (-0.32 to -0.01); 0.50 (p=0.15)	3	No data	No data	--	--
• HCC lesions <2 cm	Lesion	0.30 (0.17 to 0.43)	0.44 (0.30-0.58)	-0.14 (-0.32 to 0.05)	1	No data	No data	--	--
US without contrast (A) vs. MRI (B)	Patient	0.61 (0.48 to 0.74)	0.81 (0.69 to 0.89)	-0.19 (-0.30 to -0.08); 0.01 (p=0.79)	3	0.94 (0.87 to 0.97)	0.82 (0.66 to 0.91)	0.13 (0.03 to 0.22); 0.01 (p=0.40)	3
US without contrast (A) vs. MRI (B)	Lesion	0.57 (0.42 to 0.71)	0.79 (0.67 to 0.88)	-0.22 (-0.31 to -0.14); 0.22 (p=0.28)	3	0.75 (0.66 to 0.82)	0.78 (0.70-0.85)	-0.03 (-0.13 to 0.06); 0.001 (p=0.89)	2
• HCC lesion <2 cm	Lesion	0.40 (0.18 to 0.67)	0.65 (0.38 to 0.85)	-0.26 (-0.36 to -0.15); 0.60	2	0.71 (0.60-0.80)	0.84 (0.76 to 0.89)	-0.13 (-0.25 to -0.01); 0.006	2
US with contrast (A) vs. MRI (B)	Lesion	0.54 (0.25 to 0.80)	0.70 (0.40-0.89)	-0.16 (-0.30 to -0.02); 0.71 (p=0.31)	2	No data	No data	--	--
• HCC lesions <2 cm	Lesion	0.30 (0.17 to 0.43)	0.42 (0.28 to 0.56)	-0.12 (-0.31 to 0.07)	1	No data	No data	--	--

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
• Well-differentiated HCC Lesion	Lesion	0.43 (0.14 to 0.77)	0.36 (0.11 to 0.72)	0.07 (-0.19 to 0.33); 0.87 (p=0.34)	2	No data	No data	--	--
MRI (A) vs. CT (B)	Patient	0.88 (0.53 to 0.98)	0.82 (0.41 to 0.97)	0.06 (-0.05 to 0.17); 3.0 (p=0.21)	4	0.84 (0.70-0.92)	0.91 (0.82 to 0.96)	-0.08 (-0.16 to 0.00); 0.40 (p=0.21)	4
• Excluding high risk of bias studies	Patient	0.82 (0.75 to 0.88)	0.75 (0.68 to 0.81)	0.07 (-0.02 to 0.17); <0.0001	2	0.80 (0.57 to 0.92)	0.91 (0.77 to 0.97)	-0.11 (-0.23 to 0.01); 0.44	2
MRI (A) vs. CT (B)	Lesion	0.81 (0.77 to 0.84)	0.72 (0.67 to 0.77)	0.09 (0.06 to 0.12); 0.37 (p<0.0001)	28	0.85 (0.76 to 0.92)	0.90 (0.82 to 0.95)	-0.05 (-0.10 to 0.01); 0.43 (p=0.01)	6
• Excluding high risk of bias studies	Lesion	0.80 (0.73 to 0.85)	0.73 (0.66 to 0.79)	0.07 (0.04 to 0.10); 0.50 (p<0.0001)	19	0.87 (0.78 to 0.93)	0.93 (0.86 to 0.96)	-0.05 (-0.10 to 0.00); 0.37 (p=0.03)	5
• Non-hepatic specific contrast	Lesion	0.81 (0.74 to 0.87)	0.74 (0.65 to 0.81)	0.08 (0.04 to 0.11); 0.39 (p=0.01)	11	0.62 (0.51 to 0.72)	0.86 (0.77 to 0.93)	-0.24 (-0.37 to -0.11); <0.0001 (p=1.0)	2
• Hepatic specific contrast	Lesion	0.80 (0.74 to 0.85)	0.70 (0.62 to 0.77)	0.10 (0.06 to 0.14); 0.41 (p=0.0003)	15	0.93 (0.88 to 0.96)	0.91 (0.85 to 0.94)	0.02 (-0.03 to 0.07); 0.01 (p=0.78)	4
• HCC lesions <2 cm	Lesion	0.59 (0.43 to 0.73)	0.46 (0.32 to 0.62	0.12 (0.03 to 0.22); 0.25 (p=0.27)	3	0.84 (0.73 to 0.91)	0.80 (0.67 to 0.89)	0.04 (-0.06 to 0.14); 0.12 (p=0.41)	2
Evaluation of previously identified lesion (KQ 2)									
US without contrast (A) vs. CT (B)	Patient	0.78 (0.70-0.85)	0.89 (0.84 to 0.95)	-0.12 (-0.21 to -0.02)	1	No data	No data	--	--
US with contrast (A) vs. CT (B)	Patient	0.91 (0.85 to 0.95)	0.87 (0.79 to 0.92)	0.04 (-0.01 to 0.10); 0.17 (p=0.18)	4	0.93 (0.88 to 0.97)	0.94 (0.88 to 0.97)	-0.01 (-0.06 to 0.05); 0.07 (p=0.36)	2
• Excluding high risk of bias studies	Patient	0.90 (0.76 to 0.96)	0.84 (0.68 to 0.93)	0.06 (-0.03 to 0.15); 0.32 (p=xxxx)	2	0.94 (0.87 to 0.97)	0.94 (0.88 to 0.97)	-0.01 (-0.06 to 0.05); 0.08 (p=xxxx)	2
US with contrast (A) vs. CT (B)	Lesion	0.94 (0.89 to 0.97)	0.91 (0.85 to 0.94)	0.03 (-0.03 to 0.09); <0.0001 (p=1.0)	3	No data	No data	--	--

	Unit of Analysis	Sensitivity A (95% CI)	Sensitivity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies	Specificity A (95% CI)	Specificity B (95% CI)	Difference (95% CI); τ^2 (p value)	Number of Studies
• HCC lesion <2 cm	Lesion	0.78 (0.61 to 0.89)	0.71 (0.52 to 0.85)	0.07 (-0.01 to 0.15); 1.1 (p-0.02)	7	0.87 (0.62 to 0.97)	0.94 (0.77 to 0.98)	-0.06 (-0.15 to 0.03); 2.4 (p=0.09)	4
• Well-differentiated HCC Lesion	Lesion	0.55 (0.25 to 0.82)	0.55 (0.25 to 0.82)	0.00 (-0.30 to 0.30); 0.48 (p=0.40)	2	No data	No data	--	--
US with contrast (A) vs. MRI (B)	Patient	0.79 (0.65 to 0.94)	0.83 (0.69 to 0.97)	-0.03 (-0.24 to 0.17)	1	0.79 (0.68 to 0.90)	0.75 (0.64 to 0.87)	0.04 (-0.12 to 0.20)	1
• HCC lesion <2 cm	Patient	0.52 (0.39 to 0.64)	0.62 (0.49 to 0.74)	-0.10 (-0.27 to 0.08)	1	0.93 (0.84 to 1.0)	0.97 (0.90-1.0)	-0.03 (-0.15 to 0.08)	1
US with contrast (A) vs. MRI (B)	Lesion	0.79 (0.65 to 0.94)	0.83 (0.69 to 0.97)	-0.03 (-0.24 to 0.17)	1	0.79 (0.68 to 0.90)	0.75 (0.64 to 0.87)	0.04 (-0.12 to 0.20)	1
• HCC lesion <2 cm	Lesion	0.53 (0.28 to 0.76)	0.68 (0.43 to 0.86)	-0.16 (-0.30 to -0.02); 0.72 (p=0.25)	3	0.95 (0.85 to 0.98)	0.98 (0.91 to 0.99)	-0.03 (-0.08 to 0.02); 0.38 (p=0.43)	3
MRI (A) vs. CT (B)	Patient	0.81 (0.70-0.92)	0.74 (0.62 to 0.87)	0.06 (-0.10 to 0.23)	1	0.85 (0.72 to 0.99)	0.81 (0.66 to 0.96)	0.04 (-0.16 to 0.24)	1
MRI (A) vs. CT (B)	Lesion	0.84 (0.76 to 0.92)	0.62 (0.52 to 0.72)	0.22 (0.09 to 0.35)	1	0.36 (0.20-0.52)	0.72 (0.58 to 0.87)	-0.36 (-0.58 to -0.15)	1
Identification of metastatic HCC (KQ 4)									
PET/CT (A) vs. CT (B)	Patient (2), lesion (1)	0.82 (0.61 to 0.93)	0.85 (0.66 to 0.95)	-0.03 (-0.12 to 0.060); 0.75 (p=0.17)	3	Insufficient data	Insufficient data	--	--

CT = computed tomography; HCC = hepatocellular carcinoma; MRI =magnetic resonance imaging; PET = positron emission tomography; US = ultrasound

Table 21. Staging accuracy

Author, Year	Diagnostic Test	Dates of Imaging	Reference Standard	Country	Sample Size	Patient Population	Staging System	Stage Analysis	Correctly Staged	Over-staged	Under-staged
Baccarini, 2006[109]	CT	1996-2005	Explant	Italy	50	Liver transplant	TNM	0, 1, 2, 3, 4a	28%	20%	52%
Burrel, 2003[113]	MRI	2000-2001	Explant	Spain	50	Liver transplant	BCLC	A, B, C	59%	10%	31%
	CT								58%	4%	38%
Cheung, 2013[114]	FDG PET	2004-2010	Explant or surgical resection	China	43	Liver transplant or surgical resection	TNM	1, 2, 3	26%	NR	NR
	^{11}C-choline PET								91%	NR	NR
	Dual tracer PET								91%	NR	NR
	CT								42%	NR	NR
Freeman, 2006[56]	MRI	2003-2005	Explant	United States	285	Liver transplant	TNM	0, 1, 2, 3, 4a, 4b	40%	31%	29%
	CT				357				47%	27%	25%
	US				10				30%	30%	40%
	Two or more imaging methods				117				49%	29%	22%
Libbrecht, 2002[75]	US, CT, or MRI	2000-2001	Explant	Belgium	13	Liver transplant	NR	NR	NR	NR	NR
Lu CH, 2010[164]	CT	2006-2008	Explant	Taiwan	57	Liver transplant	NR	NR	NR	NR	NR
	MRI										
Luca, 2010[41]	CT	2004-2006	Explant	Italy	57	Liver transplant	TNM	1, 2, 3	46%	2%	52%
Ronzoni, 2007[187]	CT	2003-2006	Explant	Italy	88	Liver transplant	NR	NR	NR	NR	NR
Shah, 2006[92]	US or CT	1991-2004	Explant	Canada	118	Liver transplant	TNM	NR	NR	NR	NR
Valls, 2004[192]	CT	1995-2002	Explant	Spain	85	Liver transplant	NR	NR	NR	NR	NR
Zacherl, 2002[198]	CT	1998-2000	Explant	Austria	23	Liver transplant	TNM	1, 2, 3, 4	39%	NR	NR

BCLC = Barcelona Clinic Liver Cancer; NR = not reported; TNM = tumor nodule metastasis

Table 22. Test performance of fludeoxyglucose positron emission tomography for metastatic hepatocellular carcinoma, stratified by location of metastasis

Study, Year	Unit of Analysis	Sensitivity: Lung Metastasis	Sensitivity: Lymph Node Metastasis	Sensitivity: Bone Metastasis	Specificity: Lung Metastasis	Specificity: Lymph Node Metastasis	Specificity: Bone Metastasis
Kawaoka, 2009[a][137]	Patient	0.59 (10.7/18)	0.67 (10.7/16)	0.83 (10/12)	0.92 (14.7/16)	0.92 (16.7/18)	0.86 (20.7/24)
Kim YK, 2010 (3)[279]	Lesion	0.60 (3/5)	1.0 (3/3)	1.0 (5/5)	Not reported	Not reported	Not reported
Lee JE, 2012[156]	Patient	0.61 (14/23)	0.91 (20/22)	1.0 (11/11)	0.99[b]	0.96[b]	1.0[b]
Nagaoka, 2006[174]	Lesion	0.70 (7/10)	0.95 (21/22)	1.0 (16/16)	Not reported	Not reported	Not reported
Park JW, 2008[283]	Lesion	0.80 (16/20)	Not reported	1.0 (6/6)	Not reported	Not reported	Not reported
Sugiyama, 2004[285]	Lesion	0.42 (5/12)	1.0 (9/9)	0.80 (8/10)	Not reported	Not reported	Not reported
Wu, 2011[291]	Patient	0.80 (8/10)	Not reported	0.75 (3/4)	Not reported	Not reported	Not reported
Yoon, 2007[294]	Patient	1.0 (12/12)	1.0 (19/19)	1.0 (11/11)	0.84 (63/75)	0.94 (64/68)	1.0 (76/76)

[a] Based on average from three readers
[b] Unable to determine number of true-negatives from information provided in study

Discussion

Key Findings and Strength of Evidence

The key findings of this review, including strength of evidence grades, are summarized in Table 23. Details about factors assessed to determine the overall strength of evidence are shown in Appendix J.

The great preponderance of evidence on imaging for HCC was in the area of diagnostic test performance. However, few studies evaluated test performance of imaging for HCC in true surveillance settings of patients at high risk for HCC, but without a prior diagnosis of HCC, undergoing periodic imaging. Among the limited evidence available in this setting, there was no clear difference between US without contrast and CT, based on across-study comparisons of sensitivity. Two studies that directly compared sensitivity of US without contrast and CT did report lower sensitivity with US, but data are too limited to draw strong conclusions.[51,100]

Many more studies evaluated test performance of imaging for HCC in populations of patients undergoing treatment such as liver transplantation, hepatic resection, or ablation therapy, or in series of patients previously diagnosed with HCC or with HCC and other liver conditions. Such studies were considered as part of Key Question 1 with studies of surveillance because they were not designed to further characterize previously identified HCC lesions (the focus of Key Question 2). Rather, their purpose was to evaluate test performance for lesion identification, therefore providing information that could potentially be extrapolated to surveillance. However, we analyzed these studies separately from studies conducted in true surveillance settings, given the differences in the reason for imaging and the populations evaluated, including a generally much higher prevalence of HCC, with some studies only enrolling patients with HCC. In these studies, sensitivity was lower for US without contrast than for CT or MRI, with a difference based on within-study (direct) comparisons that ranged from 0.11 to 0.22. MRI and CT performed similarly when patients with HCC were the unit of analysis, but sensitivity of MRI was higher than CT when HCC lesions were the unit of analysis (pooled difference 0.09. 95% CI 0.06 to 12).

Ultrasound without contrast did not perform better than ultrasound with contrast for identification of HCC.[62,73] This is probably related to the short duration in which microbubble contrast is present within the liver, so that it is not possible to perform a comprehensive contrast-enhanced examination of the liver.[20] Rather, the main use of ultrasound with contrast appears to be for evaluation of previously identified focal liver lesions.

For characterization of previously identified lesions, we found no clear differences in sensitivity between US with contrast, CT, and MRI. Although some evidence was available on the accuracy of imaging modalities for distinguishing between HCC and other (non-HCC) liver lesions, it was not possible to draw strong conclusions due to variability in the types of non-HCC lesions evaluated (e.g., regenerative nodules, dysplastic nodules, hypervascular pseudolesions, hemangiomas, and others), small numbers of studies, and some inconsistency in findings.

Several factors appeared to affect estimates of test performance across different imaging modalities. Studies of patients with HCC were generally associated with somewhat higher sensitivity than studies that used HCC lesions as the unit of analysis. Studies that used explanted livers as the reference standard reported lower sensitivity than studies that used a nonexplant reference standard. Use of multiple reference standards poses a challenge to assessment of diagnostic accuracy.[297] Across imaging modalities, sensitivity was markedly lower for HCC lesions <2 cm versus those >2 cm (differences in sensitivity ranged from 0.30 to 0.39), and

further declined for lesions <10 mm in diameter. Evidence also consistently indicated substantially lower sensitivity for well-differentiated lesions than moderately- or poorly-differentiated lesions.

Evidence on the effects of other patient, tumor, and technical factors on test performance was more limited. For US, there was no clear effect of use of Doppler, lesion depth, or body mass index on test performance. For CT, some evidence indicated higher sensitivity for studies that used a contrast rate of ≥3 ml/s than those with a contrast rate <3 ml/s, and for studies that used delayed phase imaging. For MRI, hepatic-specific contrast agents were associated with slightly higher sensitivity than nonspecific contrast agents, but there were no clear effects of magnetic field strength (3.0 vs. 1.5 T), use of delayed phase imaging, timing of delayed phase imaging ≥120 seconds after administration of contrast of <120 s), section thickness (≤5 mm vs. >5 mm), or use of diffusion-weighted imaging. For identification of intrahepatic HCC lesions, limited evidence found PET with ^{11}C-acetate and other alternative tracers such as ^{18}F-fluorocholine and ^{18}F-fluorothymidine associated with substantially higher sensitivity than FDG PET. Sensitivity of FDG PET was lower than sensitivity of FDG PET/CT.

Few studies evaluated the comparative test performance of multiple imaging modalities, either in combination or sequentially as part of a diagnostic algorithm. The limited available evidence suggests that using multiple imaging tests and defining a positive test as typical imaging findings on at least one imaging modality increases sensitivity without substantively reducing specificity.

Conclusions were generally robust on sensitivity, and we stratified analyses based on study factors such as setting (Asia vs. United States or Europe), prospective collection of data, interpretation of imaging findings blinded to results of the reference standard, avoidance of case-control design, and overall risk of bias.

Across analyses, specificity was generally high, with most pooled estimates around 0.85 or higher, and few clear differences between imaging modalities. However, many studies did not report specificity and pooled estimates of specificity were frequently imprecise, precluding strong conclusions regarding comparative test performance. Since likelihood ratios are sensitive to small changes in estimates when the specificity is high, it was also difficult to draw strong conclusions regarding comparative diagnostic test performance based on differences in positive or negative likelihood ratios. Most likelihood ratio estimates fell into or near the "moderately useful" range (positive likelihood ratio of 5 to 10 and negative likelihood ratio of 0.1 to 0.2), with the exception of FDG PET for identification of intrahepatic HCC lesions, which was associated with a negative likelihood ratio of 0.50.

Evidence regarding the accuracy of imaging modalities for staging was primarily limited to CT. Most studies addressed accuracy of CT, with 28 percent to 58 percent correctly staged based on TNM criteria, with somewhat more understaging (25% to 52%) than overstaging (2% to 27%). Studies on the accuracy of imaging for identifying metastatic HCC disease were primarily limited to FDG PET or PET/CT, with a pooled sensitivity of 0.82 to 0.85.

Evidence on the comparative effectiveness of imaging for HCC on diagnostic thinking, use of subsequent procedures, or resource utilization was extremely limited. In studies that compared the accuracy of transplant decisions based on CT against primarily explanted livers as the reference standard, the proportion correctly assessed for transplant eligibility based on Milan criteria ranged from 40 percent to 96 percent. Evidence on the effects of surveillance with imaging versus no surveillance on clinical outcomes was limited to a single randomized trial.[47] Although it found an association between surveillance with US and AFP and decreased liver-

specific mortality, the trial was conducted in China, potentially limiting applicability to screening to the United States, and had important methodological shortcomings.

Evidence on comparative harms associated with imaging was also extremely limited, with no study measuring downstream harms related to false-positive tests or subsequent workup, or potential harms related to labeling or psychological effects. A handful of studies reported low rates of serious direct harms (e.g., allergic reactions) associated with imaging. However, evidence on administration of contrast for radiological procedures in general also suggest a low rate of serious adverse events. For example, a retrospective analysis of over 450,000 doses of low-osmolar iodinated or gadolinium contrast administered at a single center identified a total of 522 adverse events (0.11% of total), with the most frequent adverse events being urticaria (52%) and nausea (18%).[298] Fewer than 100 of the events required further treatment, with use of epinephrine in nine instances. The rate of adverse events was 0.15 percent for iodinated contrast and 0.04 percent for gadolinium, consistent with estimates from other studies.[299-301]

No study on US with contrast reported harms. Potential harms associated with use of microbubble contrast agents were highlighted when the FDA issued a black box warning in 2007 regarding use of perflutren microbubble contrast for cardiac imaging, due to reports of four fatalities due to cardiopulmonary events within 30 minutes of perflutren administration and 11 fatalities within 12 hours.[302] Other studies have attempted to quantify rates of harms associated with microbubble contrast. One study of sulfur hexafluoride contrast for various abdominal applications (23,188 imaging studies) reported 29 adverse events, with two rated serious (0.01%); there were no deaths.[303] A study of 16,025 patients who received perflutren contrast in cardiac imaging reported an overall adverse event rate of 0.12 percent, with a rate of serious adverse events of 0.04 percent and no deaths.[304]

Although PET and CT are associated with risk of radiation exposure, no study of imaging for HCC was designed to evaluate potential long-term clinical outcomes associated with radiation exposure. According to the Radiological Society of North American and the American College of Radiology, abdominal CT with and without contrast is associated with an approximate effective radiation dose of 20 mSv and PET/CT with 25 mSV.[305]

Findings in Relationship to What is Already Known

Unlike our review, several previously published reviews on detection of HCC and evaluation of focal liver lesions found no clear differences in test performance between US, CT, and MRI for HCC.[306-309] Several factors may explain these discrepancies—we included more studies than any prior review, separately analyzed studies based on the reason for imaging, stratified studies according to the unit of analysis, and focused on within-study (direct) comparisons of two or more imaging modalities against a common reference standard instead of relying primarily or solely on across-study (indirect) estimates of test performance. Research on meta-analyses of diagnostic tests found that conclusions based on such direct comparisons are often different from conclusions based on indirect comparisons, and may therefore be more suitable for comparing diagnostic tests.[36] In fact, a recently published meta-analysis that focused on direct comparisons was consistent with our review in finding MRI with hepatic-specific contrast associated with higher sensitivity than CT.[310] Our review is consistent with previous reviews regarding lower sensitivity of imaging for detection of small and well-differentiated HCC lesions.

Our findings regarding test performance of PET for detection of metastatic HCC are consistent with a recently published systematic review and meta-analysis that reported a pooled sensitivity of 0.77.[311] Like our review, a recent systematic review found insufficient evidence to

determine effects of surveillance with imaging on clinical outcomes.[312] A systematic review on screening for HCC in chronic liver disease funded by the U.S. Department of Veterans Affairs is currently in progress.[32]

Applicability

A number of potential issues could impact the applicability of our findings. Over half of the studies were conducted in Asia, where the prevalence, underlying causes, course, evaluation, and management of chronic liver disease may be different than in the United States. To mitigate potential effects of study country on applicability, we excluded invasive imaging techniques not typically used in the United States such as CT arterial portography and CT hepatic arteriography, as well as imaging techniques considered inadequate in the United States (such as C-arm CT). We also performed stratified analyses focusing on studies performed in the United States and Europe to evaluate effects on estimates of diagnostic accuracy and found no clear effects on estimates.

Imaging techniques are rapidly evolving, which is another factor that could affect applicability. To mitigate effects of outdated techniques on applicability, we excluded imaging technologies considered outdated, such as MRI with magnetic field strength <1.5 T and nonspiral CT, and only included studies published since 1998. We also performed additional analyses on technical factors such as contrast rate, imaging phases evaluated, timing of imaging phases, section thickness, use of hepatobiliary contrast (for MRI), use of diffusion-weighted imaging, and newer technologies such as dual-source or spectral CT. We included studies of US with microbubble contrast event though no agent is currently approved for abdominal imaging in the United States, because efforts to obtain FDA approval are ongoing and this technique is commonly used in other geographic areas of the world, including Canada and Europe.

As noted above, few studies were performed in true surveillance settings, i.e., in patients at high risk for HCC but not previously diagnosed with this condition. Rather, most studies of test performance that were not performed specifically to evaluate or characterize previously identified lesions were conducted in patients undergoing imaging for other reasons, including series of patients undergoing liver transplantation, surgical resection, or other treatments for HCC. Although such studies are likely to provide some useful findings regarding diagnostic accuracy, results may not be directly applicable to patients undergoing surveillance. In particular, the high prevalence of HCC (many studies only enrolled patients with HCC) could overestimate test performance in true surveillance settings, in which the prevalence of HCC would be much lower.[313]

Implications for Clinical and Policy Decisionmaking

Our review has important potential implications for clinical and policy decisionmaking. Due to the lack of direct evidence regarding clinical benefits and downstream harms associated with different imaging tests for surveillance, diagnosis, and staging of HCC, most decisions regarding use of imaging tests must necessarily be made primarily on the basis of diagnostic test performance. Despite limited evidence in true surveillance settings, our study support current recommendations from the AASLD for US without contrast for surveillance of HCC in at-risk populations.[9] Although sensitivity of CT and MRI for identifying HCC was higher than US in studies conducted in nonsurveillance settings, findings may not be directly applicable to clinical and policy decisions related to surveillance, as the spectrum of patients evaluated in these studies could have affected estimates.

In patients found to have an HCC lesion on surveillance, our review supports use of CT and MRI to further characterize lesions >1 cm in size, as in the AASLD guideline, based on high sensitivity and specificity. Evidence is very limited but appears consistent with the sequential diagnostic imaging algorithm as outlined in the AASLD guideline, in which typical findings for HCC on sequentially performed CT or MRI are considered sufficient to make a diagnosis.

Our findings also support minimal technical specifications for MRI and CT for HCC imaging as suggested in recent guidance, such as those regarding minimum contrast rates and use of delayed phase imaging.[15] Evidence suggesting superior test performance of MRI with hepatic-specific versus nonhepatic contrast appears promising, though differences were relatively small. Therefore, clinical and policy decisions around use of nonhepatic contrast may be impacted by additional factors other than test performance, such as cost, harms, or convenience. For example, maximum increase in liver parenchyma signal intensity with hepatic-specific contrast agents is achieved after 20 minutes to hours following contrast administration.[314]

Although US with contrast was associated with similar test performance as MRI and CT for evaluation of lesions, no microbubble contrast agents are currently approved for use in the United States. Although the role of PET is likely to remain focused on identification of metastatic HCC and staging, additional research could help clarify the role of PET with alternative tracers for identification and evaluation of intrahepatic HCC.

Clinicians and policymakers may consider modeling studies to help estimate potential benefits and harms of screening. For models to appropriately inform decision making, however, requires reliable estimates of important input parameters such as subsequent testing, interventions, and associated benefits and harms that occur as a result of imaging. Such data are not currently available.

Limitations of the Review Process

Substantial statistical heterogeneity was present in most pooled analyses of diagnostic accuracy; this situation is common in meta-analyses of diagnostic accuracy.[315-317] As noted in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, "heterogeneity is to be expected in meta-analyses of diagnostic test accuracy."[317] To address the anticipated heterogeneity, we utilized random effects models to pool studies and stratified studies according to the reason that imaging was performed and the unit of analysis used. We also performed additional stratified and sensitivity analyses based on the reference standard used, study characteristics (such as country in which the study was conducted, factors related to risk of bias), patient characteristics, and technical factors related to the imaging tests under investigation. As noted previously, results were generally robust in sensitivity analyses, despite the heterogeneity. Due to the relatively small numbers of studies, we were unable to perform meaningful meta-regression. We also focused on evaluations of comparative test performance based on within-study comparisons of imaging modalities, which tended to be associated with less heterogeneity than pooled across-study estimates. However, a limitation of our analysis of within-group comparisons is that we had to treat the two compared groups as independent, because we had only aggregated data. Individual patient level data would be required to take into account the paired nature of the comparisons.

We were unable to construct a summary receiver operating characteristic curve, because most studies did not use a ratings scale to classify imaging tests as positive or negative, and the scales that were used differed across studies (e.g., 1-3, 0-4, 1-4, 1-5, and others). We also did not attempt to pool summary measures of discrimination, for several reasons. Some studies reported

the area under the receiver operating characteristic (AUROC) curve and others reported the alternate free response operating characteristic (AFROC) curve, and the suitability of pooling such measures is uncertain. In addition, a number of studies that reported the AUROC or AFROC did not report specificity, and it was unclear from the data provided in the studies how the measures were calculated. Finally, it was often unclear whether the AUROC or AFROC was constructed based on different cutoffs for sensitivity and specificity (representing a true area under a curve) or based on a single cutoff for sensitivity and specificity.

We excluded non-English language articles and did not search for studies published only as abstracts. We did not formally assess for publication bias using statistical or graphical methods for assessing sample size effects, as research indicates that such methods can be seriously misleading.[318,319] Although we found no evidence of unpublished studies through searches on clinical trial registries and regulatory documents, the usefulness of such methods for identifying unpublished studies of test performance is likely to be limited.

Limitations of the Evidence Base

We identified a number of limitations of the evidence base on imaging for HCC. Only one clinical trial with important methodological shortcomings has evaluated clinical outcomes associated with surveillance for HCC in high-risk patients versus no screening,[47] and no trial has compared effects of different imaging modalities for screening. Evidence on effects of imaging on diagnostic thinking, subsequent procedures, and resource utilization is also extremely sparse. There is almost no evidence comparing harms associated with different imaging modalities or strategies.

Despite identifying over 200 studies on test performance, we also found important limitations related to these outcomes. Only three studies were rated low risk of bias and 86 studies were rated high risk of bias. Nearly half of the studies did not avoid use of a case-control design, which can result in spectrum bias and inflated estimates of diagnostic accuracy. In addition, nearly half of the studies did not clearly report interpretation of imaging findings blinded to the results of the reference standard test. Many studies did not report specificity, particularly for lesion-based analyses of diagnostic accuracy, perhaps due to the difficulty in defining a "true negative" lesion in such situations. Estimates for pooled specificity were therefore incomplete and typically imprecise, as were likelihood ratio estimates, which are calculated from pooled sensitivity and specificity.

Other limitations include relatively limited numbers of direct comparisons of diagnostic accuracy between different imaging modalities and techniques (i.e., studies that perform two or more imaging techniques in the same population and evaluate diagnostic accuracy of each technique against the same reference standard). Research has shown that results from such direct comparisons are often different from results based on indirect comparisons (i.e., comparisons of different tests in across studies performed in different populations). Therefore, we focused on results from direct comparisons when possible.

We were unable to evaluate a number of potentially important technical factors in the studies, such as the type of contrast injector and use of bolus-tracking methods for CT; type of contrast injector, use of bolus-tracking methods, spatial resolution, and length of breath hold for MRI; and effects of reader experience and training and transducer frequency for US. Evidence for newer techniques such as spectral or dual-source CT was also limited to only a few studies. For evaluation of the effects of patient and tumor characteristics on measures of diagnostic accuracy, most of the evidence focused on effects of tumor size and degree of differentiation, with very

little evidence on patient characteristics such as age, race, sex, severity of liver disease, or underlying cause of liver disease.

Research Gaps

Significant research gaps limit the full understanding of the comparative effectiveness of imaging for surveillance, diagnosis, and staging of HCC. The only randomized trial of effects of surveillance for HCC with imaging on clinical outcomes had important methodological shortcomings and was performed in China, potentially limiting applicability to screening in the United States[47] Although conducting a randomized trial of surveillance versus no screening in the United States could be difficult because screening is recommended in clinical practice guidelines and routinely performed in high-risk patients, randomized trials that compare screening using different imaging modalities or combinations of modalities would be helpful for understanding optimal approaches. In particular, studies assessing clinical outcomes associated with application of the AASLD algorithm versus alternative strategies would be very informative. Potential challenges in conducting such studies include the need to enroll large samples with sufficient statistical power and with lengthy followup.

In lieu of such studies, evidence on effects of alternative imaging strategies on intermediate outcomes such as diagnostic thinking, subsequent procedures, and resource utilization could also be informative. Such studies could potentially enroll smaller samples and would probably not require the extended followup needed to assess clinical outcomes.

Although many studies are available on test performance of alternative imaging modalities and strategies, important research gaps remain. Notably, few studies evaluated imaging in true surveillance settings, and evidence on accuracy of imaging for identifying HCC lesions from nonsurveillance settings may not be directly applicable to surveillance due to spectrum effects. More studies are also needed to clarify the role of promising alternative techniques, such as US with contrast, MRI with hepatic-specific contrast, and PET with alternative tracers, on estimates of accuracy. Research should focus on improving methods for identifying small or well-differentiated HCC lesions, for which imaging remains suboptimal.

To be most informative it is important for studies to utilize methods for reducing bias in the conduct of studies of test performance, such as avoidance of case-control design and use of methods to insure interpretation of imaging tests blinded to results of the reference standard. Another important shortcoming of the available literature is the failure of many studies to report specificity, resulting in incomplete and less precise estimates. Given the difficulty in defining true negatives for studies that use HCC lesions as the unit of analysis, we suggest that investigators consider routinely reporting results using patients as the unit of analysis, though HCC lesion-based analyses may be reported in addition. Finally, additional studies that evaluate different imaging modalities, techniques, or strategies against a common reference standard in the same population would be more helpful for understanding comparative test performance than studies that evaluate a single imaging modality or technique.

Conclusions

Based on estimates of test performance, several imaging modalities appear to be reasonable options for surveillance, diagnosis, or staging of HCC. Although there are some potential differences in test performance between different imaging modalities and techniques, more research is needed to understand the effects of such differences on diagnostic thinking and clinical outcomes.

Table 23. Summary of evidence on imaging techniques for the surveillance, diagnosis, and staging of hepatocellular carcinoma

Key Question 1. Surveillance

Key Question 1a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Surveillance settings *Unit of analysis: patients with HCC*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.82 (95% CI 0.66 to 0.092, 3 studies) and specificity 0.87 (95% CI 0.77 to 0.93, 2 studies), for a LR+ of 6.2 (95% CI 3.6 to 11) and LR- of 0.20 (0.10 to 0.40).
	CT	Sensitivity: Low Specificity: Low	Sensitivity was 0.84 (95% CI 0.59 to 0.95, 2 studies) and specificity 0.99 (95% CI 0.86 to 0.999, 2 studies).
	MRI or PET	Insufficient	No evidence
Surveillance settings *Unit of analysis: HCC lesions*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.60 (95% CI 0.36 to 0l.80, 2 studies) and specificity 0.94 (95% CI 0.83 to 0.98, 1 study), for a LR+ of 9.8 (95% CI 3.7 to 26) and LR- of 0.43 (95% CI 0.24 to 0.74.
	CT	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.62 (95% CI 0.46 to 0.76, 1 study).
	MRI or PET	No evidence	No evidence
Nonsurveillance settings *Unit of analysis: patients with HCC*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.73 (95% CI 0.46 to 0.90, 8 studies) and specificity 0.93 (95% CI 0.85 to 0.97, 6 studies), for a LR+ of 11 (95% CI 5.4 to 21) and LR- of 0.29 (95% CI 0.13 to 0.65).
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.83 (95% CI 0.75 to 0.89, 16 studies) and specificity 0.92 (95% CI 0.86 to 0.96, 11 studies), for a LR+ of 11 (95% CI 5.6 to 20) and LR- of 0.19 (95% CI 0.12 to 0.28).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.87 (95% CI 0.77 to 0.93, 11 studies) and specificity 0.88 (95% CI 0.79 to 0.93, 9 studies), for a LR+ of 7.2 (95% CI 3.9 to 13) and LR- of 0.15 (95% CI 0.08 to 0.27).
	PET	Sensitivity: Low Specificity: Low	FDG PET: Sensitivity was 0.52 (95% CI 0.39 to 0.66, 15 studies) and specificity 0.95 (95% CI 0.92 to 0.99, 5 studies), for a LR+ of 11 (95% CI 2.6 to 49) and LR- of 0.50 (95% CI 0.37 to 0.68). ^{11}C-acetate PET: Sensitivity was 0.85 (95% CI 0.67 to 0.94, 4 studies). Specificity was not reported.
Nonsurveillance settings *Unit of analysis: HCC lesions*	US without contrast	Sensitivity: Low Specificity: Low	Sensitivity was 0.60 (95% CI 0.42 to 0.75, 11 studies). Only 2 studies reported specificity, with inconsistent results (0.63, 95% CI 0.53 to 0.73 and 0.95, 95% CI 0.85 to 0.99).
	US with contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.76 (95% CI 0.53 to 0.90, 6 studies). No study evaluated specificity.
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.77 (95% CI 0.73 to 0.81, 75 studies) and specificity 0.89 (95% CI 0.83 to 0.93, 20 studies), for a LR+ of 7.0 (95% CI 4.6 to 11) and LR- of 0.25 (95% CI 0.21 to 0.30).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.83 (95% CI 0.80 to 0.86, 69 studies) and specificity 0.83 (95% CI 0.70 to 0.92, 13 studies), for a LR+ of 5.0 and LR- of 0.20 (95% CI 0.16 to 0.2.6).
	PET	Sensitivity: Low Specificity: Low	FDG PET: Sensitivity was 0.56 (95% CI 0.41 to 0.69, 4 studies) and specificity 0.91 (95% CI 0.76 to 0.98, 1 study). ^{11}C-acetate PET: Sensitivity was 0.78 (95% CI 0.61 to

	Imaging Modality or Comparison	Strength of Evidence	Summary
			0.89, 4 studies). Specificity was not reported.
Direct (within-study) comparisons of imaging Modalities *Unit of analysis: Patients with HCC*	US without contrast vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.68 (95% CI 0.54 to 0.80) vs. 0.80 (95% CI 0.68 to 0.88), for a difference of -0.12 (95% CI -0.20 to -0.03), based on 6 studies.
	US without contrast vs. MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.61 (95% CI 0.48 to 0.74) vs. 0.81 (95% CI 0.69 to 0.89), for a difference of -0.19 (95% CI -0.30 to -0.08), based on 3 studies.
	MRI vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.88 (95% CI 0.53 to 0.98) vs. 0.82 (95% CI 0.41 to 0.97), for a difference of 0.06 (95% CI -0.05 to 0.17), based on 4 studies.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: HCC lesions*	US without contrast vs. CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.55 (95% CI 0.43 to 0.66) vs. 0.66 (95% CI 0.54 to 0.76) for a difference of -0.11 (95% CI -0.18 to -0.04), based on 3 studies.
	US without contrast vs. MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.57 (95% CI 0.42 to 0.71) vs. 0.79 (95% CI 0.67 to 0.88), for a difference of -0.22 (95% -0.31 to 0.14), based on 3 studies.
	US with contrast vs. CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.58 (95% CI 0.37 to 0.77 vs. 0.74 (95% CI 0.54 to 0.87), for a difference of -0.16 (95% CI -0.32 to 0.01), based on 3 studies.
	US with contrast vs. MRI	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.25 to 0.80) vs. 0.70 (95% CI 0.40 to 0.89), for a difference of -0.16 (95% CI -0.30 to -0.02), based on 2 studies.
	MRI vs. CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.81 (95% CI 0.77 to 0.84) vs. 0.72 (95% CI 0.67 to 0.77), for a difference of 0.09 (95% CI 0.06 to 0.12), based on 28 studies. Findings were similar when studies were stratified according to use of non-hepatic specific or hepatic specific contrast.
Multiple imaging modalities	Various combinations	Sensitivity: Low	1 study found sensitivity of imaging with various combinations of two imaging modalities was similar or lower than single modality imaging, based on concordant positive findings on 2 imaging modalities. The other study reported higher sensitivity with multiple imaging modalities than with single modality imaging, but criteria for positive results based on multiple imaging modalities were unclear

Key Question 1a.i. Effects of reference standard on test performance (based on HCC lesions as the unit of analysis)

Imaging Modality or Comparison	Strength of Evidence	Summary
US	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.34 (95% CI 0.21 to 0.49) in 5 studies that used explanted liver as the reference standard and ranged from 0.70 to 0.85 in studies that used other reference standards.
CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.69 (95% CI 0.60 to 0.77) in 21 studies that used explanted liver as the reference standard and ranged from 0.79 to 0.85 in studies that used other reference standards.
MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.69 (95% CI 0.59 to 0.77) in 15 studies that used explanted liver as the reference standard and ranged from 0.85 to 0.88 in studies that used other reference standards.
PET	Sensitivity: Low Specificity: Insufficient	No study of FDG PET used an explanted liver reference standard.

Key Question 1a.ii. Effects of patient, tumor, technical, and other factors on test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Lesion size	US	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.78 to 0.94) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions <2 cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies. The difference was larger in studies of US without contrast than studies of US with contrast, but these findings are difficult to interpret because sensitivity for HCC lesions <20 mm was much lower in the studies of US without contrast. For US without contrast, sensitivity was 0.09 (95% CI 0.02 to 0.29, 4 studies) for lesions < 10 mm to 0.50 (95% CI 0.23 to 0.78, 4 studies) for lesions 10 to 20 mm and 0.88 (95% CI 0.66 to 0.96, 4 studies) for lesions >20 mm , for a difference of 0.37 (95% CI 0.18 to 0.57) for lesions >20 mm vs. 10 to 20 mm, and 0.41 (95% CI 0.19 to 0.63) for lesions 10 to 20 mm vs. <10 mm. For ultrasound with contrast, three studies found sensitivity of 0.64 (95% CI 0.33 to 0.87) for lesions 10 to 20 mm and 0.91 (95% CI 0.71 to 0.98) for lesions >20 mm, for a difference of 0.26 (95% CI 0.04 to 0.48).
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.91 to 0.95) for lesions >2 cm and 0.62 (95% CI 0.56 to 0.68) for lesions <2 cm, for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36), based on 33 studies. Sensitivity was 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10 to 20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 vs. 10 to 20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10 to 20 vs. <10 mm.
	MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.96 (95% CI 0.93 to 0.97) for lesions >2 cm and 0.65 (95% CI 0.57 to 0.73) for lesions <2 cm, for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37), based on 25 studies. Sensitivity was 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10 to 20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 vs. 10 to 20 mm and 0.34 (95% CI 0.27 to 0.41) for 10 to 20 vs. <10 mm.
	PET	Sensitivity: Low	For FDG PET, sensitivity was consistently higher for larger lesions, based on 5 studies. Data were not pooled due to differences in the tumor size categories evaluated. Two studies of ^{11}C-acetatate PET found inconsistent effects of lesion size on sensitivity
Degree of tumor differentiation	US with contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.83 (95% CI 0.55 to 0.95) for moderately or poorly-differentiated HCC lesions and 0.43 (95% CI 0.15 to 0.76) for well differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.17 to 0.64), based on 3 studies.
	CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.82 (95% CI 0.66 to 0.91) for moderately or poorly-differentiated HCC lesions and 0.50 (95% CI 0.29 to 0.70) for well differentiated lesions, for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45), based on 5 studies.
	MRI	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.26 to 0.79) for moderately or poorly-differentiated HCC lesions and 0.38 (95% CI 0.17 to 0.64) for well differentiated lesions, but the difference was not statistically significant (0.16, 95% CI -0.11 to 0.43), based on 2 studies.

	Imaging Modality or Comparison	Strength of Evidence	Summary
	PET	Sensitivity: Low Specificity: Insufficient	For FDG PET, sensitivity was consistently higher for more poorly-differentiated lesions than more well differentiated lesions, based on five studies. In three studies of ^{11}C-acetate PET and one study of ^{18}F-fluorochlorine, sensitivity for more well differentiated lesions was not lower than more poorly-differentiated lesions.
Other factors	US	Low	In 2 studies that directly compared US with versus without contrast, there was no clear difference in sensitivity (-0.04, 95% CI -0.11 to 0.04). 1 study that directly compared use of Doppler versus no Doppler showed no clear effect on estimates of sensitivity. Lesion depth and body mass index had no effect on estimates of sensitivity.
	CT	Low-Moderate	Using patients with HCC as the unit of analysis, studies with a contrast rate ≥3 ml/s reported a higher sensitivity (0.87, 95% CI 0.77 to 0.93, 8 studies) than studies with a contrast rate <3 ml/s (0.71, 95% CI 0.50 to 0.85, 4 studies) and studies with delayed phase imaging reported somewhat higher sensitivity (0.89, 95% CI 0.81 to 0.94, 7 studies) than studies without delayed phase imaging (0.74, 95% CI 0.66 to 0.87, 7 studies), but there were no clear effects in studies that used HCC lesions as the unit of analysis.
	MRI	Low-Moderate	There were no clear differences in estimates of diagnostic accuracy when studies were stratified according to MRI scanner type (1.5 vs. 3.0 T), imaging phases evaluated (with or without delayed phase imaging), timing of delayed phase imaging (>120 seconds vs. <120 seconds), section thickness (≤5 mm for enhanced images vs. >5 mm), or use of diffusion-weighted imaging. In studies that directly compared diagnostic accuracy with different types of contrast, hepatic-specific contrast agents were associated with slightly higher sensitivity than non-hepatic specific contrast agents (0.82, 95% CI 0.71 to 0.90 vs. 0.75, 95% CI 0.61 to 0.85, difference 0.07, 95% CI 0.01 to 0.14, 5 studies).
	PET	Low-Moderate	FDG PET was associated with lower sensitivity that ^{11}C-acetate PET when either patients (0.58 vs. 0.81, for a difference of -0.23, 95% CI -0.34 to -0.13, 3 studies**)** or HCC lesions (0.52 vs. 0.79, for a difference of -0.27, 95 %CI -0.36 to -0.17, 3 studies) were the unit of analysis. FDG PET was also associated with lower sensitivity that dual tracer PET with FDG and ^{11}C-acetate or ^{18}F-choline PET, but evidence was limited to 1 or 2 studies for each of these comparisons. Using patients as the unit of analysis, sensitivity of FDG PET (0.39, 95% CI 0.24 to 0.56, 8 studies) was lower than sensitivity of FDG PET/CT (0.65, 95% CI 0.50 to 0.78, 7 studies).

Key Question 1b. Diagnostic thinking

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 1c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
US plus serum AFP	Low	1 cluster randomized controlled trial (n=18816) conducted in China found screening every 6 months with noncontrast US plus serum AFP versus no screening in persons 35 to 79 years of age (mean 42 years)

		with HBV infection or chronic hepatitis without HBV infection associated with lower risk of HCC-related mortality (32 vs. 54 deaths, rate ratio 0.63, 95% CI 0.41 to 0.98) at 5 year followup, but was rated high risk of bias. 2 trials found no clear differences in mortality with US screening at 4- vs. 12-month intervals, or at 3- vs. 6-month intervals.

Key Question 1d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
MRI, CT, US	Insufficient	1 study reported no serious adverse events associated with administration of gadoxetic acid for MRI and one study reported no clear differences in adverse events between CT with contrast at 3 ml/s versus 5 ml/s. No study reported rates of adverse events associated with use of microbubble contrast agents in US, and harms were not reported in randomized trials of screening with imaging.

Key Question 2. Diagnosis

Key Question 2a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Evaluation of a previously identified lesion *Unit of analysis: Patients with HCC*	US with contrast	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.79 to 0.94, 8 studies) and specificity 0.92 (95% CI 0.84 to 0.96, 5 studies), for a LR+ of 11 (95% CI 5.5 to 20) and LR- of 0.13 (95% CI 0.07 to 0.24).
	US without contrast	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.78 (95% CI 0.72 to 0.83) in 2 studies; specificity was not reported.
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.85 (95% CI 0.76 to 0.91, 5 studies) and specificity 0.92 (95% CI 0.86 to 0.96, 3 studies), for a LR+ of 11 (95% CI 5.7 to 22) and LR- of 0.17 (95% CI 0.10 to 0.27).
	MRI	Sensitivity: Low Specificity: Low	Sensitivity was 0.76 (95% CI 0.62 to 0.86, 3 studies) and specificity 0.87 (95% CI 0.70 to 0.95, 3 studies), for a LR+ of 5.9 (95% CI 2.5 to 14) and LR- of 0.28 (95% CI 0.18 to 0.43).
Evaluation of a previously identified lesion *Unit of analysis: HCC lesions*	US with contrast	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.86 (95% CI 0.79 to 0.91, 21 studies) and specificity 0.93 (95% CI 0.87 to 0.96, 11 studies) for a LR+ of 12 (95% CI 6.3 to 21) and LR- of 0.15 (95% CI 0.10 to 0.23).
	CT	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.80 (95% CI 0.67 to 0.88, 12 studies) and specificity 0.89 (95% CI 0.29 to 0.99, 6 studies), for a LR+ of 6.9 (95% CI 0.53 to 91) and LR- of 0.23 (95% CI 0.13 to 0.40).
	MRI	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.79 (95% CI 0.69 to 0.87, 13 studies) and specificity 0.95 (95% CI 0.82 to 0.99, 12 studies), for a LR+ of 15 (95% CI 4.4 to 50) and LR- of 0.22 (95% CI 0.15 to 0.33).
	PET	Sensitivity: Moderate Specificity: Moderate	Sensitivity was 0.56 to 0.57 and specificity 1.0 in 2 studies of FDG PET.
For distinguishing HCC lesions from non-HCC hepatic lesions	US with contrast	Low	1 study found US with sulfur hexafluoride contrast associated with a sensitivity of 0.94 (62/66) and specificity of 0.68 (23/34) for distinguishing hypervascular HCC from focal nodular hyperplasia, using quantitative methods.
	CT	Low	4 studies evaluated accuracy of CT for distinguishing HCC from non-HCC lesions, but the non-HCC lesions varied in the studies, precluding strong conclusions.

	Imaging Modality or Comparison	Strength of Evidence	Summary
	MRI	Moderate	4 studies reported inconsistent results for distinguishing small (<2 to 3 cm) hypervascular HCC lesions from hypervascular pseudolesions, with sensitivity 0.47 and 0.52 in 2 studies, and 0.91 and 0.92 in the other two. Specificity was 0.93 or higher in all four studies. Five other studies evaluated accuracy of MRI for distinguishing HCC from other non-HCC lesions, but the non-HCC lesions varied in the studies.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: Patients with HCC*	US without contrast vs. CT	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.78 (95% CI 0.70 to 0.85) vs. 0.89 (95% CI 0.84 to 0.95), for a difference of -0.12 (95% CI -0.21 to -0.02), based on 1 study.
	US with contrast vs. CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.91 (0.85 to 0.95) vs. 0.87 (95% CI 0.79 to 0.92), for a difference of 0.04 (95% CI -0.01 to 0.10), based on 4 studies.
	MRI vs. CT	Sensitivity: Low Specificity: Low	Sensitivity was 0.81 (95% CI 0.70 to 0.92) vs. 0.74 (95% CI 0.62 to 0.87), for a difference of 0.06 (-0.10 to 0.23), based on 1 study.
Direct (within-study) comparisons of imaging modalities *Unit of analysis: HCC lesion*	US with contrast vs. CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.89 to 0.97) vs. 0.91 (95% CI 0.85 to 0.94), for a difference of 0.03 (95% CI -0.03 to 0.09), based on 3 studies.
	US with contrast vs. MRI	Sensitivity: Low Specificity: Low	Sensitivity was 0.79 (95% CI 0.65 to 0.94) vs. 0.83 (95% CI 0.69 to 0.97), for a difference of -0.03 (95% CI -0.24 to 0.17), based on 1 study.
	MRI vs. CT	Sensitivity: Low Specificity: Low	One study found MRI associated with higher sensitivity (0.84, 95% CI 0.76 to 0.92 vs. 0.62, 95% CI 0.52 to 0.72, for a difference of 0.22, 95% CI 0.09 to 0.35) but lower specificity (0.36, 95% CI 0.20 to 0.52 vs. 0.72, 95% CI 0.58 to 0.87, for a difference of -0.36, 95% CI -0.58 to -0.15) than CT.
Multiple imaging modalities	Various combinations	Moderate	In 4 studies in which positive results with multiple modality imaging were defined as concordant typical findings for HCC on 2 imaging modalities, sensitivity was lower than with a single modality (difference in sensitivity ranged from 0.09 to 0.27), with no clear difference in specificity. In three studies in which positive results with multiple modality imaging were defined as typical findings for HCC on at least one of the imaging techniques, sensitivity was higher than with a single modality (increase in sensitivity ranged from 0.09 to 0.25), with no clear difference in specificity. 1 study found that a sequential imaging strategy in which a second imaging test was only performed for indeterminant results on initial CT increased sensitivity for HCC from 0.53 to 0.74 to 0.79.

Key Question 2a.i. Effects of reference standard on test performance (based on HCC lesions as the unit of analysis)

Imaging Modality or Comparison	Strength of Evidence	Summary
All	Sensitivity: Moderate Specificity: Moderate	No study used explanted liver as the reference standard. There were no clear differences across imaging modalities in estimates of diagnostic accuracy in analyses stratified by use of different non-explant reference standards.

Key Question 2a.ii. Effects of patient, tumor, technical, and other factors on test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Lesion size	US	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.88 (95% CI 0.78 to 0.94) for lesions >2 cm and 0.49 (95% CI 0.31 to 0.67) for lesions > 2cm, for an absolute difference in sensitivity of 0.39 (95% CI 0.27 to 0.51), based on 14 studies.
	CT	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.94 (95% CI 0.91 to 0.95) for lesions >2 cm and 0.62 (95% CI 0.56 to 0.68) for lesions <2 cm, for an absolute difference in sensitivity of 0.31 (95% CI 0.26 to 0.36), based on 33 studies. Sensitivity was 0.32 (95% CI 0.24 to 0.40, 20 studies) for lesions <10 mm, 0.73 (95% CI 0.66 to 0.80, 22 studies) for lesions 10 to 20 mm, and 0.95 (95% CI 0.92 to 0.97, 19 studies), for a difference of 0.42 (95% CI 0.35 to 0.48) for lesions >20 vs. 10 to 20 mm and 0.21 (95% CI 0.15 to 0.27) for lesions 10 to 20 vs. <10 mm.
	MRI	Sensitivity: Moderate Specificity: Low	Sensitivity was 0.96 (95% CI 0.93 to 0.97) for lesions >2 cm and 0.65 (95% CI 0.57 to 0.73) for lesions <2 cm, for an absolute difference in sensitivity of 0.30 (95% CI 0.23 to 0.37), based on 25 studies. Sensitivity was 0.43 (95% CI 0.32 to 0.54, 19 studies) for lesions <10 mm, 0.77 (95% CI 0.67 to 0.84, 18 studies) for lesions 10 to 20 mm, and 0.97 (95% CI 0.94 to 0.98, 14 studies) for lesions >20 mm (0.97, 95% CI 0.94 to 0.98), for a difference of 0.20 (95% CI 0.13 to 0.28) for >20 vs. 10 to 20 mm and 0.34 (95% CI 0.27 to 0.41) for 10 to 20 vs. <10 mm.
Degree of tumor differentiation	US	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.84 (95% CI 0.64 to 0.94) for moderately or poorly-differentiated HCC lesions and 0.43 (95% CI 0.21 to 0.69) for well differentiated lesions, for an absolute difference in sensitivity of 0.40 (95% CI 0.22 to 0.59), based on 4 studies.
	CT	Sensitivity: Moderate Specificity: Insufficient	Sensitivity was 0.82 (95% CI 0.66 to 0.91) for moderately or poorly-differentiated HCC lesions and 0.50 (95% CI 0.29 to 0.70) for well differentiated lesions, for an absolute difference in sensitivity of 0.32 (95% CI 0.19 to 0.45), based on 5 studies.
	MRI	Sensitivity: Low Specificity: Insufficient	Sensitivity was 0.54 (95% CI 0.26 to 0.79) for moderately or poorly-differentiated HCC lesions and 0.38 (95% CI 0.17 to 0.64) for well differentiated lesions, but the difference was not statistically significant (0.16, 95% CI - 0.11 to 0.43), based on 2 studies.
Other factors	US	Insufficient-Low	In 2 studies that directly compared US with versus without contrast, US with contrast was associated with sensitivity of 0.89 (95% CI 0.83 to 0.93) and US without contrast with a sensitivity of (0.39) 95% CI 0.32 to 0.47), for a difference in sensitivity of 0.50 (95% CI 0.41 to 0.58). Based on across-study comparisons, there were no clear differences in sensitivity between different US contrast agents; no study directly compared different contrast agents. There were no differences in sensitivity of US based on lesion depth (3 studies) or body mass index (2 studies).
	CT	Insufficient-Low	Evidence on effects of technical parameters (type of CT scanner, use of delayed phase imaging, section thickness) was limited by small numbers of studies with wide confidence intervals and methodological limitations, precluding reliable conclusions. 2 studies found no clear difference in sensitivity of CT for HCC in patients with versus without cirrhosis.
	MRI	Low-Moderate	There were no clear differences in estimates of sensitivity based on the type of MRI machine (3.0 T vs. 1.5 T), type of

			contrast, use of delayed phase imaging, timing of delayed phase imaging, and section thickness. Estimates were similar were studies that used diffusion-weighted imaging were excluded.

Key Question 2b. Diagnostic thinking

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 2c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

Key Question 2d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
US and CT	Insufficient	1 study of US (with and without contrast) and CT reported harms, but did not stratify results by imaging technique. The overall rate of adverse drug-related events was 10%, with all events classified as mild.

Key Question 3. Staging

Key Question 3a. Test performance

	Imaging Modality or Comparison	Strength of Evidence	Summary
Staging accuracy, using TNM criteria	CT	Moderate	The proportion correctly staged ranged from 28% to 58%, the proportion overstaged from 2% to 27%, and the proportion understaged from 25% to 52%, based on 6 studies.
	MRI	Low	The proportion correctly staged were 10% and 31%, the proportion overstaged 10% and 31%, and the proportion understaged 29% and 31%, based on 2 studies.
	PET	Low	1 study found 26% of patients were correctly staged with FDG PET and 91% with ^{11}C-choline PET.
	MRI vs. CT	Low	2 studies reported similar staging accuracy.
Identification of metastatic disease *Unit of analysis: Patients with metastatic HCC*	PET	Sensitivity: Moderate Specificity: Moderate	Sensitivity of FDG PET was 0.85 (95% CI 0.71 to 0.93, 6 studies) and specificity 0.93 (95% CI 0.89 to 0.95, 5 studies), for a LR+ of 11 (95% CI 7.8 to17) and LR- of 0.16 (95% CI 0.08 to 0.33). 1 study that directly compared sensitivity of FDG PET to 11-chloroacetate PET reported comparable sensitivity (0.79 vs. 0.71), though sensitivity was higher when both tracers were used (0.98).
Identification of metastatic disease *Unit of analysis: Metastatic HCC lesions*	PET	Sensitivity: Moderate Specificity: Insufficient	Sensitivity of FDG PET was 0.82 (95% CI 0.72 to 0.90, 5 studies). 1 study that directly compared sensitivity of FDG to 11-choloroacetate PET reported comparable sensitivity (0.86 vs. 0.77, respectively).

Key Question 3.a.i. Effects of reference standard on test performance

Imaging Modality or Comparison	Strength of Evidence	Summary
CT, MRI, PET	Sensitivity: Low Specificity: Low	Evidence was insufficient to determine effect s of different reference standards on accuracy of staging using TNM criteria or accuracy of PET for identifying metastatic HCC because few studies evaluated alternative reference standards.

Key Question 3.a.ii. Effects of patient, tumor, and technical factors on test performance

Imaging Modality or Comparison	Strength of Evidence	Summary
CT, MRI, PET	No evidence	For accuracy of staging using TNM criteria, no study evaluated effects of patient-level characteristics or other factors on accuracy of imaging techniques for staging.
PET	Low-Moderate	In 1 study that directly compared sensitivity of PET vs. PET/CT for identifying metastatic HCC lesions, there was no clear difference in sensitivity. 4 studies of FDG PET found sensitivity increased as lesion size increased, but the number of lesions <1 cm was small (total of 20). 8 studies generally found sensitivity of FDG PET higher for lymph and bone metastasis than for lung metastasis, but samples were small, precluding strong conclusions.

Key Question 3b. Diagnostic thinking

	Imaging Modality or Comparison	Strength of Evidence	Summary
Transplant eligibility, using Milan criteria	CT	Moderate	The proportion correctly assessed for transplant eligibility ranged from 40% to 96%. The proportion of patients who met transplant criteria based on CT but exceeded criteria based on the reference standard was 3.5 to 7.8%, based on 3 studies. 2 studies found that 2.3% and 16% of patients who underwent transplantation based on Milan criteria had no HCC lesions on examination of explanted livers.
	CT vs. MRI	Low	1 study reported similar accuracy.
	PET vs. CT	Low	1 study found ^{11}C-choline PET more accurate than CT (95% vs. 40%).
Use of resection and ablative therapies	MRI vs. CT	Low	1 study reported that the proportion of decisions to perform resection or ablative therapies that were classified as correct were similar for MRI (90% and 90%, respectively) and CT (80% and 77%, respectively).

Key Question 3c. Clinical and patient-centered outcomes

Imaging Modality or Comparison	Strength of Evidence	Summary
US with contrast vs. US without contrast plus CT	Low	1 cohort study found that contrast enhanced US identified more small (≤2 cm) HCC lesions than noncontrast US plus CT (36 vs. 31), and was associated with a higher complete necrosis rate following ablation (92% or 106/115 vs. 83% or 93/112 lesions, p=0.036), but was rated high risk of bias.

Key Question 3d. Harms

Imaging Modality or Comparison	Strength of Evidence	Summary
All	No evidence	No evidence

AFP = alpha-fetoprotein; CT = computed tomography; FDG = fludeoxyglucose; HBV = hepatitis B virus; HCC = hepatocellular carcinoma; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; MRI = magnetic resonance imaging; PET = positron emission tomography; TNM = tumor, node, metastasis staging; US = ultrasound

References

1. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol. 2001 Sep;2(9):533-43. PMID: 11905707.
2. National Cancer Institute. Liver Cancer. http://www.cancer.gov/cancertopics/types/liver. Accessed on September 21 2011.
3. Howlader NN, A; Neyman, N; Altekruse, SF; Kosary, CL; Yu, M; Ruhl, J; Tatalovich, Z; Cho, H; Mariotto, A; Lewis, DR; Chen, HS; Feuer, EJ; Cronin, KA, ed SEER Cancer Statistics Review, 1975-2010. Bethesda, MD: National Cancer Institute; 2013.
4. Kohler BA, Ward E, McCarthy BJ, et al. Annual report to the nation on the status of cancer, 1975-2007, featuring tumors of the brain and other nervous system. J Natl Cancer Inst. 2011 May 4;103(9):714-36. PMID: 21454908.
5. McGlynn KA, London WT. Epidemiology and natural history of hepatocellular carcinoma. Best Pract Res Clin Gastroenterol. 2005 Feb;19(1):3-23. PMID: 15757802.
6. El-Serag HB. Epidemiology of hepatocellular carcinoma in USA. Hepatol Res. 2007 Sep;37 Suppl 2:S88-94. PMID: 17877502.
7. El-Serag HB. Hepatocellular carcinoma: an epidemiologic view. J Clin Gastroenterol. 2002 Nov-Dec;35(5 Suppl 2):S72-8. PMID: 12394209.
8. Zanetti AR, Van Damme P, Shouval D. The global impact of vaccination against hepatitis B: a historical overview. Vaccine. 2008 Nov 18;26(49):6266-73. PMID: 18848855.
9. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology. 2011 Mar;53(3):1020-2. PMID: 21374666.
10. Cabibbo G, Enea M, Attanasio M, et al. A meta-analysis of survival rates of untreated patients in randomized clinical trials of hepatocellular carcinoma. Hepatology. 2010 Apr;51(4):1274-83. PMID: 20112254.
11. Pons F, Varela M, Llovet JM. Staging systems in hepatocellular carcinoma. HPB (Oxford). 2005;7(1):35-41.
12. Forner A, Reig ME, de Lope CR, et al. Current strategy for staging and treatment: the BCLC update and future prospects. Semin Liver Dis. 2010 Feb;30(1):61-74. PMID: 20175034.
13. Mazzaferro V, Bhoori S, Sposito C, et al. Milan criteria in liver transplantation for hepatocellular carcinoma: An evidence-based analysis of 15 years of experience. Liver Transpl. 2011;17(S2):S44-S57.
14. Wilson SR, Greenbaum LD, Goldberg BB. Contrast-enhanced ultrasound: what is the evidence and what are the obstacles? AJR Am J Roentgenol. 2009 Jul;193(1):55-60. PMID: 19542395.
15. Wald C, Russo MW, Heimbach JK, et al. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology. 2013 Feb;266(2):376-82. PMID: 23362092.
16. Seale MK, Catalano OA, Saini S, et al. Hepatobiliary-specific MR Contrast Agents: Role in Imaging the Liver and Biliary Tree. Radiographics. 2009 October 1, 2009;29(6):1725-48.
17. Frydrychowicz A, Lubner MG, Brown JJ, et al. Hepatobiliary MR imaging with gadolinium-based contrast agents. J Magn Reson Imaging. 2012 Mar;35(3):492-511. PMID: 22334493.
18. Altenbernd J, Heusner TA, Ringelstein A, et al. Dual-energy-CT of hypervascular liver lesions in patients with HCC: investigation of image quality and sensitivity. Eur Radiol. 2011 Apr;21(4):738-43. PMID: 20936520.
19. Silva AC, Morse BG, Hara AK, et al. Dual-Energy (Spectral) CT: Applications in Abdominal Imaging. Radiographics. 2011 July 1, 2011;31(4):1031-46.
20. Lencioni R, Piscaglia F, Bolondi L. Contrast-enhanced ultrasound in the diagnosis of hepatocellular carcinoma. J Hepatol. 2008 May;48(5):848-57. PMID: 18328590.
21. Treglia G, Giovannini E, Di Franco D, et al. The role of positron emission tomography using carbon-11 and fluorine-18 choline in tumors other than prostate cancer: a systematic review. Ann Nucl Med. 2012 Jul;26(6):451-61. PMID: 22566040.
22. Li C, Zhang Z, Zhang P, et al. Diagnostic accuracy of des-gamma-carboxy prothrombin versus alpha-fetoprotein for hepatocellular carcinoma: A systematic review. Hepatol Res. 2013 Jul 9PMID: 23834468.
23. Di Tommaso L, Franchi G, Park YN, et al. Diagnostic value of HSP70, glypican 3, and glutamine synthetase in hepatocellular nodules in cirrhosis. Hepatology. 2007 Mar;45(3):725-34. PMID: 17326147.
24. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(13)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2013. Chapters available at: www.effectivehealthcare.ahrq.gov.

25. Chang SM, Matchar DB, eds. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2012. Available at http://effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.

26. Imaging Techniques for the Surveillance, Diagnosis, and Staging of Hepatocellular Carcinoma. Research Protocol. Rockville, MD: Agency for Healthcare Research and Quality; 2013. http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1600&pageaction=displayproduct. Accessed on November 11, 2013.

27. Imaging techniques for the surveillance, diagnosis, and staging of hepatocellular carcinoma. PROSPERO 2013:CRD42013005246 2013. Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013005246. Accessed on November 11, 2013.

28. Submit Scientific Information Packets. Rockville, MD: Agency for Healthcare Research and Quality; 2013. http://effectivehealthcare.ahrq.gov/index.cfm/submit-scientific-information-packets/. Accessed on November 12, 2013.

29. Barr RG. Off-label use of ultrasound contrast agents for abdominal imaging in the United States. Journal of Ultrasound in Medicine. 2013 Jan;32(1):7-12. PMID: 23269705.

30. Greenbaum L, Burns P, Copel J, et al. American Institute of Ultrasound in Medicine Recommendations for Contrast-Enhanced Liver Ultrasound Imaging Clinical Trials. Journal of Ultrasound in Medicine. 2007 June 1, 2007;26(6):705-16.

31. Nelson TR, Fowlkes JB. Contrast-Enhanced Ultrasound: An Idea Whose Time Has Come. Journal of Ultrasound in Medicine. 2007 June 1, 2007;26(6):703-4.

32. ESP Reports in Progress. Washington, DC: U.S. Department of Veterans Affairs. Health Services Research & Development Services.; 2013. http://www.hsrd.research.va.gov/publications/esp/in_progress.cfm. Accessed on November 11, 2013.

33. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001 Apr;20(3 Suppl):21-35. PMID: 11306229.

34. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36. PMID: 22007046.

35. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(13)-EHC063-EF. Chapter 9. Rockville, MD: Agency for Healthcare Research and Quality; 2013. Available at http://www.effectivehealthcare.ahrq.gov/.

36. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013 Apr 2;158(7):544-54. PMID: 23546566.

37. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol. 2006 Dec;59(12):1331-2; author reply 2-3. PMID: 17098577.

38. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. Stat Med. 2008 Feb 28;27(5):687-97. PMID: 17611957.

39. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". J Gen Intern Med. 2012 Jun;27 Suppl 1:S56-66. PMID: 22648676.

40. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1198-207. PMID: 21463926.

41. Luca A, Caruso S, Milazzo M, et al. Multidetector-row computed tomography (MDCT) for the diagnosis of hepatocellular carcinoma in cirrhotic candidates for liver transplantation: prevalence of radiological vascular patterns and histological correlation with liver explants.[Erratum appears in Eur Radiol. 2011 Jul;21(7):1574 Note: Grutttadauria, Salvatore [corrected to Gruttadauria, Salvatore]]. Eur Radiol. 2010 Apr;20(4):898-907. PMID: 19802612.

42. Sangiovanni A, Manini MA, Iavarone M, et al. The diagnostic and economic impact of contrast imaging techniques in the diagnosis of small hepatocellular carcinoma in cirrhosis. Gut. 2010 May;59(5):638-44. PMID: 19951909.

43. Serste T, Barrau V, Ozenne V, et al. Accuracy and disagreement of computed tomography and magnetic resonance imaging for the diagnosis of small hepatocellular carcinoma and dysplastic nodules: role of biopsy. Hepatology. 2012 Mar;55(3):800-6. PMID: 22006503.

44. Chen MH, Wu W, Yang W, et al. The use of contrast-enhanced ultrasonography in the selection of patients with hepatocellular carcinoma for radio frequency ablation therapy. Journal of Ultrasound in Medicine. 2007;26(8):1055-63. PMID: 17646367.

45. Trinchet JC, Chaffaut C, Bourcier V, et al. Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3- and 6-month periodicities. Hepatology. 2011 Dec;54(6):1987-97. PMID: 22144108.

46. Wang JH, Chang KC, Kee KM, et al. Hepatocellular carcinoma surveillance at 4- vs. 12-month intervals for patients with chronic viral hepatitis: a randomized study in community. Am J Gastroenterol. 2013 Mar;108(3):416-24. PMID: 23318478.

47. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004 Jul;130(7):417-22. PMID: 15042359.

48. Alaboudy A, Inoue T, Hatanaka K, et al. Usefulness of combination of imaging modalities in the diagnosis of hepatocellular carcinoma using Sonazoid(R)-enhanced ultrasound, gadolinium diethylene-triamine-pentaacetic acid-enhanced magnetic resonance imaging, and contrast-enhanced computed tomography. Oncology. 2011;81 Suppl 1:66-72. PMID: 22212939.

49. Bennett GL, Krinsky GA, Abitbol RJ, et al. Sonographic detection of hepatocellular carcinoma and dysplastic nodules in cirrhosis: correlation of pretransplantation sonography and liver explant pathology in 200 patients. AJR Am J Roentgenol. 2002 Jul;179(1):75-80. PMID: 12076908.

50. Catala V, Nicolau C, Vilana R, et al. Characterization of focal liver lesions: comparative study of contrast-enhanced ultrasound versus spiral computed tomography. Eur Radiol. 2007 Apr;17(4):1066-73. PMID: 17072617.

51. Chalasani N, Horlander JC, Sr., Said A, et al. Screening for hepatocellular carcinoma in patients with advanced cirrhosis. Am J Gastroenterol. 1999 Oct;94(10):2988-93. PMID: 10520857.

52. Dai Y, Chen MH, Fan ZH, et al. Diagnosis of small hepatic nodules detected by surveillance ultrasound in patients with cirrhosis: Comparison between contrast-enhanced ultrasound and contrast-enhanced helical computed tomography. Hepatol Res. 2008 Mar;38(3):281-90. PMID: 17908168.

53. Di Martino M, Marin D, Guerrisi A, et al. Intraindividual comparison of gadoxetate disodium-enhanced MR imaging and 64-section multidetector CT in the Detection of hepatocellular carcinoma in patients with cirrhosis. Radiology. 2010 Sep;256(3):806-16. PMID: 20720069.

54. D'Onofrio M, Rozzanigo U, Masinielli BM, et al. Hypoechoic focal liver lesions: characterization with contrast enhanced ultrasonography. J Clin Ultrasound. 2005 May;33(4):164-72. PMID: 15856516.

55. Forner A, Vilana R, Ayuso C, et al. Diagnosis of hepatic nodules 20 mm or smaller in cirrhosis: Prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma.[Erratum appears in Hepatology. 2008 Feb;47(2):769]. Hepatology. 2008 Jan;47(1):97-104. PMID: 18069697.

56. Freeman RB, Mithoefer A, Ruthazer R, et al. Optimizing staging for hepatocellular carcinoma before liver transplantation: A retrospective analysis of the UNOS/OPTN database. Liver Transpl. 2006 Oct;12(10):1504-11. PMID: 16952174.

57. Furuse J, Maru Y, Yoshino M, et al. Assessment of arterial tumor vascularity in small hepatocellular carcinoma. Comparison between color doppler ultrasonography and radiographic imagings with contrast medium: dynamic CT, angiography, and CT hepatic arteriography. Eur J Radiol. 2000 Oct;36(1):20-7. PMID: 10996754.

58. Gaiani S, Celli N, Piscaglia F, et al. Usefulness of contrast-enhanced perfusional sonography in the assessment of hepatocellular carcinoma hypervascular at spiral computed tomography. J Hepatol. 2004 Sep;41(3):421-6. PMID: 15336445.

59. Gambarin-Gelwan M, Wolf DC, Shapiro R, et al. Sensitivity of commonly available screening tests in detecting hepatocellular carcinoma in cirrhotic patients undergoing liver transplantation. Am J Gastroenterol. 2000 Jun;95(6):1535-8. PMID: 10894592.

60. Giorgio A, De Stefano G, Coppola C, et al. Contrast-enhanced sonography in the characterization of small hepatocellular carcinomas in cirrhotic patients: comparison with contrast-enhanced ultrafast magnetic resonance imaging. Anticancer Res. 2007 Nov-Dec;27(6C):4263-9. PMID: 18214030.

61. Giorgio A, Ferraioli G, Tarantino L, et al. Contrast-enhanced sonographic appearance of hepatocellular carcinoma in patients with cirrhosis: Comparison with contrast-enhanced helical CT appearance. AJR Am J Roentgenol. 2004;183(5):1319-26. PMID: 15505297.

62. Goto E, Masuzaki R, Tateishi R, et al. Value of post-vascular phase (Kupffer imaging) by contrast-enhanced ultrasonography using Sonazoid in the detection of hepatocellular carcinoma. J Gastroenterol. 2012;47(4):477-85. PMID: 22200940.

63. Hatanaka K, Kudo M, Minami Y, et al. Differential diagnosis of hepatic tumors: value of contrast-enhanced harmonic sonography using the newly developed contrast agent, Sonazoid. Intervirology. 2008;51 Suppl 1:61-9. PMID: 18544950.

64. Iavarone M, Sangiovanni A, Forzenigo LV, et al. Diagnosis of hepatocellular carcinoma in cirrhosis by dynamic contrast imaging: the importance of

tumor cell differentiation. Hepatology. 2010 Nov;52(5):1723-30. PMID: 20842697.

65. Imamura M, Shiratori Y, Shiina S, et al. Power Doppler sonography for hepatocellular carcinoma: factors affecting the power Doppler signals of the tumors. Liver. 1998 Dec;18(6):427-33. PMID: 9869398.

66. Inoue T, Kudo M, Hatanaka K, et al. Imaging of hepatocellular carcinoma: Qualitative and quantitative analysis of postvascular phase contrast-enhanced ultrasonography with sonazoid. Oncology. 2008;75(SUPPL. 1):48-54. PMID: 19092272.

67. Inoue T, Kudo M, Maenishi O, et al. Value of liver parenchymal phase contrast-enhanced sonography to diagnose premalignant and borderline lesions and overt hepatocellular carcinoma. AJR Am J Roentgenol. 2009 Mar;192(3):698-705. PMID: 19234266.

68. Kawada N, Ohkawa K, Tanaka S, et al. Improved diagnosis of well-differentiated hepatocellular carcinoma with gadolinium ethoxybenzyl diethylene triamine pentaacetic acid-enhanced magnetic resonance imaging and Sonazoid contrast-enhanced ultrasonography. Hepatol Res. 2010;40(9):930-6. PMID: 20887598.

69. Khalili K, Kim TK, Jang H-J, et al. Optimization of imaging diagnosis of 1-2 cm hepatocellular carcinoma: an analysis of diagnostic performance and resource utilization. J Hepatol. 2011 Apr;54(4):723-8. PMID: 21156219.

70. Kim CK, Lim JH, Lee WJ. Detection of hepatocellular carcinomas and dysplastic nodules in cirrhotic liver: accuracy of ultrasonography in transplant patients. Journal of Ultrasound in Medicine. 2001 Feb;20(2):99-104. PMID: 11211142.

71. Kim PN, Choi D, Rhim H, et al. Planning ultrasound for percutaneous radiofrequency ablation to treat small (<= 3 cm) hepatocellular carcinomas detected on computed tomography or magnetic resonance imaging: a multicenter prospective study to assess factors affecting ultrasound visibility. J Vasc Interv Radiol. 2012 May;23(5):627-34. PMID: 22387030.

72. Korenaga K, Korenaga M, Furukawa M, et al. Usefulness of Sonazoid contrast-enhanced ultrasonography for hepatocellular carcinoma: Comparison with pathological diagnosis and superparamagnetic iron oxide magnetic resonance images. J Gastroenterol. 2009;44(7):733-41. PMID: 19387532.

73. Kunishi Y, Numata K, Morimoto M, et al. Efficacy of fusion imaging combining sonography and hepatobiliary phase MRI with Gd-EOB-DTPA to detect small hepatocellular carcinoma. AJR Am J Roentgenol. 2012 Jan;198(1):106-14. PMID: 22194485.

74. Lee MW, Kim YJ, Park HS, et al. Targeted sonography for small hepatocellular carcinoma discovered by CT or MRI: factors affecting sonographic detection. AJR Am J Roentgenol. 2010 May;194(5):W396-400. PMID: 20410384.

75. Libbrecht L, Bielen D, Verslype C, et al. Focal lesions in cirrhotic explant livers: pathological evaluation and accuracy of pretransplantation imaging examinations. Liver Transpl. 2002 Sep;8(9):749-61. PMID: 12200773.

76. Lim JH, Kim SH, Lee WJ, et al. Ultrasonographic detection of hepatocellular carcinoma: Correlation of preoperative ultrasonography and resected liver pathology. Clin Radiol. 2006;61(2):191-7. PMID: 16439225.

77. Liu WC, Lim JH, Park CK, et al. Poor sensitivity of sonography in detection of hepatocellular carcinoma in advanced liver cirrhosis: accuracy of pretransplantation sonography in 118 patients. Eur Radiol. 2003 Jul;13(7):1693-8. PMID: 12835987.

78. Luo W, Numata K, Kondo M, et al. Sonazoid-enhanced ultrasonography for evaluation of the enhancement patterns of focal liver tumors in the late phase by intermittent imaging with a high mechanical index. Journal of Ultrasound in Medicine. 2009 Apr;28(4):439-48. PMID: 19321671.

79. Luo W, Numata K, Morimoto M, et al. Focal liver tumors: characterization with 3D perflubutane microbubble contrast agent-enhanced US versus 3D contrast-enhanced multidetector CT. Radiology. 2009 Apr;251(1):287-95. PMID: 19221060.

80. Luo W, Numata K, Morimoto M, et al. Three-dimensional contrast-enhanced sonography of vascular patterns of focal liver tumors: pilot study of visualization methods. AJR Am J Roentgenol. 2009 Jan;192(1):165-73. PMID: 19098197.

81. Mita K, Kim SR, Kudo M, et al. Diagnostic sensitivity of imaging modalities for hepatocellular carcinoma smaller than 2 cm. World J Gastroenterol. 2010 Sep 7;16(33):4187-92. PMID: 20806437.

82. Mok TSK, Yu SCH, Lee C, et al. False-negative rate of abdominal sonography for detecting hepatocellular carcinoma in patients with hepatitis B and elevated serum alpha-fetoprotein levels. AJR Am J Roentgenol. 2004 Aug;183(2):453-8. PMID: 15269040.

83. Moriyasu F, Itoh K. Efficacy of perflubutane microbubble-enhanced ultrasound in the characterization and detection of focal liver lesions: phase 3 multicenter clinical trial. AJR Am J Roentgenol. 2009 Jul;193(1):86-95. PMID: 19542399.

84. Ooi C-C, Low S-CA, Schneider-Kolsky M, et al. Diagnostic accuracy of contrast-enhanced ultrasound in differentiating benign and malignant

focal liver lesions: a retrospective study. J Med Imaging Radiat Oncol. 2010 Oct;54(5):421-30. PMID: 20958940.

85. Paul SB, Gulati MS, Sreenivas V, et al. Evaluating patients with cirrhosis for hepatocellular carcinoma: value of clinical symptomatology, imaging and alpha-fetoprotein. Oncology. 2007;72 Suppl 1:117-23. PMID: 18087192.

86. Pei XQ, Liu LZ, Xiong YH, et al. Quantitative analysis of contrast-enhanced ultrasonography: Differentiating focal nodular hyperplasia from hepatocellular carcinoma. Br J Radiol. 2013 Mar;86(1023):20120536 Epub 2013 Feb 7. PMID: 23392189.

87. Quaia E, Alaimo V, Baratella E, et al. The added diagnostic value of 64-row multidetector CT combined with contrast-enhanced US in the evaluation of hepatocellular nodule vascularity: implications in the diagnosis of malignancy in patients with liver cirrhosis. Eur Radiol. 2009 Mar;19(3):651-63. PMID: 18815790.

88. Rickes S, Schulze S, Neye H, et al. Improved diagnosing of small hepatocellular carcinomas by echo-enhanced power Doppler sonography in patients with cirrhosis. Eur J Gastroenterol Hepatol. 2003 Aug;15(8):893-900. PMID: 12867800.

89. Rode A, Bancel B, Douek P, et al. Small nodule detection in cirrhotic livers: evaluation with US, spiral CT, and MRI and correlation with pathologic examination of explanted liver. J Comput Assist Tomogr. 2001 May-Jun;25(3):327-36. PMID: 11351179.

90. Seitz K, Bernatik T, Strobel D, et al. Contrast-enhanced ultrasound (CEUS) for the characterization of focal liver lesions in clinical practice (DEGUM Multicenter Trial): CEUS vs. MRI--a prospective comparison in 269 patients. Ultraschall Med. 2010 Oct;31(5):492-9. PMID: 20652854.

91. Seitz K, Strobel D, Bernatik T, et al. Contrast-Enhanced Ultrasound (CEUS) for the characterization of focal liver lesions - prospective comparison in clinical practice: CEUS vs. CT (DEGUM multicenter trial). Parts of this manuscript were presented at the Ultrasound Dreilandertreffen 2008, Davos. Ultraschall Med. 2009 Aug;30(4):383-9. PMID: 19688670.

92. Shah SA, Tan JC, McGilvray ID, et al. Accuracy of staging as a predictor for recurrence after liver transplantation for hepatocellular carcinoma. Transplantation. 2006 Jun 27;81(12):1633-9. PMID: 16794527.

93. Singh P, Erickson RA, Mukhopadhyay P, et al. EUS for detection of the hepatocellular carcinoma: results of a prospective study. Gastrointest Endosc. 2007 Aug;66(2):265-73. PMID: 17543307.

94. Strobel D, Raeker S, Martus P, et al. Phase inversion harmonic imaging versus contrast-enhanced power Doppler sonography for the characterization of focal liver lesions. Int J Colorectal Dis. 2003 Jan;18(1):63-72. PMID: 12458384.

95. Sugimoto K, Moriyasu F, Saito K, et al. Comparison of Kupffer-phase Sonazoid-enhanced sonography and hepatobiliary-phase gadoxetic acid-enhanced magnetic resonance imaging of hepatocellular carcinoma and correlation with histologic grading. Journal of Ultrasound in Medicine. 2012 Apr;31(4):529-38. PMID: 22441909.

96. Suzuki S, Iijima H, Moriyasu F, et al. Differential diagnosis of hepatic nodules using delayed parenchymal phase imaging of levovist contrast ultrasound: comparative study with SPIO-MRI. Hepatol Res. 2004 Jun;29(2):122-6. PMID: 15163434.

97. Tanaka S, Ioka T, Oshikawa O, et al. Dynamic sonography of hepatic tumors. AJR Am J Roentgenol. 2001 Oct;177(4):799-805. PMID: 11566675.

98. Teefey SA, Hildeboldt CC, Dehdashti F, et al. Detection of primary hepatic malignancy in liver transplant candidates: prospective comparison of CT, MR imaging, US, and PET. Radiology. 2003 Feb;226(2):533-42. PMID: 12563151.

99. Trojan J, Schroeder O, Raedle J, et al. Fluorine-18 FDG positron emission tomography for imaging of hepatocellular carcinoma. Am J Gastroenterol. 1999 Nov;94(11):3314-9. PMID: 10566736.

100. Van Thiel DH, Yong S, Li SD, et al. The development of de novo hepatocellular carcinoma in patients on a liver transplant list: frequency, size, and assessment of current screening methods. Liver Transpl. 2004 May;10(5):631-7. PMID: 15108254.

101. Wang ZL, Tang J, Weskott HP, et al. Undetermined focal liver lesions on gray-scale ultrasound in patients with fatty liver: characterization with contrast-enhanced ultrasound. J Gastroenterol Hepatol. 2008 Oct;23(10):1511-9. PMID: 18713302.

102. Xu HX, Lu MD, Liu LN, et al. Discrimination between neoplastic and non-neoplastic lesions in cirrhotic liver using contrast-enhanced ultrasound. Br J Radiol. 2012 Oct;85(1018):1376-84. PMID: 22553290.

103. Xu H-X, Xie X-Y, Lu M-D, et al. Contrast-enhanced sonography in the diagnosis of small hepatocellular carcinoma < or =2 cm. J Clin Ultrasound. 2008 Jun;36(5):257-66. PMID: 18088056.

104. Yamamoto K, Shiraki K, Deguchi M, et al. Diagnosis of hepatocellular carcinoma using digital subtraction imaging with the contrast agent, Levovist: comparison with helical CT, digital subtraction angiography, and US angiography.

Oncol Rep. 2002 Jul-Aug;9(4):789-92. PMID: 12066210.

105. Yu NC, Chaudhari V, Raman SS, et al. CT and MRI improve detection of hepatocellular carcinoma, compared with ultrasound alone, in patients with cirrhosis. Clin Gastroenterol Hepatol. 2011 Feb;9(2):161-7. PMID: 20920597.

106. Zhou K-R, Yan F-H, Tu B-W. Arterial phase of biphase enhancement spiral CT in diagnosis of small hepatocellular carcinoma. Hepatobiliary Pancreat Dis Int. 2002 Feb;1(1):68-71. PMID: 14607626.

107. Addley HC, Griffin N, Shaw AS, et al. Accuracy of hepatocellular carcinoma detection on multidetector CT in a transplant liver population with explant liver correlation. Clin Radiol. 2011 Apr;66(4):349-56. PMID: 21295772.

108. Akai H, Kiryu S, Matsuda I, et al. Detection of hepatocellular carcinoma by Gd-EOB-DTPA-enhanced liver MRI: comparison with triple phase 64 detector row helical CT. Eur J Radiol. 2011 Nov;80(2):310-5. PMID: 20732773.

109. Baccarani U, Adani GL, Avellini C, et al. Comparison of clinical and pathological staging and long-term results of liver transplantation for hepatocellular carcinoma in a single transplant center. Transplant Proc. 2006 May;38(4):1111-3. PMID: 16757280.

110. Baek CK, Choi JY, Kim KA, et al. Hepatocellular carcinoma in patients with chronic liver disease: a comparison of gadoxetic acid-enhanced MRI and multiphasic MDCT. Clin Radiol. 2012 Feb;67(2):148-56. PMID: 21920517.

111. Bartolozzi C, Donati F, Cioni D, et al. MnDPDP-enhanced MRI vs dual-phase spiral CT in the detection of hepatocellular carcinoma in cirrhosis. Eur Radiol. 2000;10(11):1697-702. PMID: 11097390.

112. Bhattacharjya S, Bhattacharjya T, Quaglia A, et al. Liver transplantation in cirrhotic patients with small hepatocellular carcinoma: an analysis of pre-operative imaging, explant histology and prognostic histologic indicators. Dig Surg. 2004;21(2):152-9; discussion 9-60. PMID: 15166485.

113. Burrel M, Llovet JM, Ayuso C, et al. MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation. Hepatology. 2003 Oct;38(4):1034-42. PMID: 14512891.

114. Cheung TT, Ho CL, Lo CM, et al. 11C-acetate and 18F-FDG PET/CT for clinical staging and selection of patients with hepatocellular carcinoma for liver transplantation on the basis of milan criteria: Surgeon's perspective. J Nucl Med. 2013;54(2):192-200. PMID: 23321459.

115. Colagrande S, Fargnoli R, Dal Pozzo F, et al. Value of hepatic arterial phase CT versus lipiodol ultrafluid CT in the detection of hepatocellular carcinoma. J Comput Assist Tomogr. 2000 Nov-Dec;24(6):878-83. PMID: 11105704.

116. de Ledinghen V, Laharie D, Lecesne R, et al. Detection of nodules in liver cirrhosis: spiral computed tomography or magnetic resonance imaging? A prospective study of 88 nodules in 34 patients. Eur J Gastroenterol Hepatol. 2002 Feb;14(2):159-65. PMID: 11981340.

117. Denecke T, Grieser C, Froling V, et al. Multislice computed tomography using a triple-phase contrast protocol for preoperative assessment of hepatic tumor load in patients with hepatocellular carcinoma before liver transplantation. Transpl Int. 2009 Apr;22(4):395-402. PMID: 19000231.

118. Di Martino M, De Filippis G, De Santis A, et al. Hepatocellular carcinoma in cirrhotic patients: prospective comparison of US, CT and MR imaging. Eur Radiol. 2013;23:887-96. PMID: 23179521.

119. Doyle DJ, O'Malley ME, Jang H-J, et al. Value of the unenhanced phase for detection of hepatocellular carcinomas 3 cm or less when performing multiphase computed tomography in patients with cirrhosis. J Comput Assist Tomogr. 2007 Jan-Feb;31(1):86-92. PMID: 17259838.

120. Freeny PC, Grossholz M, Kaakaji K, et al. Significance of hyperattenuating and contrast-enhancing hepatic nodules detected in the cirrhotic liver during arterial phase helical CT in pre-liver transplant patients: radiologic-histopathologic correlation of explanted livers. Abdom Imaging. 2003 May-Jun;28(3):333-46. PMID: 12719903.

121. Golfieri R, Marini E, Bazzocchi A, et al. Small (<or=3 cm) hepatocellular carcinoma in cirrhosis: the role of double contrast agents in MR imaging vs. multidetector-row CT. Radiol Med (Torino). 2009 Dec;114(8):1239-66. PMID: 19697104.

122. Haradome H, Grazioli L, Tinti R, et al. Additional value of gadoxetic acid-DTPA-enhanced hepatobiliary phase MR imaging in the diagnosis of early-stage hepatocellular carcinoma: comparison with dynamic triple-phase multidetector CT imaging. J Magn Reson Imaging. 2011 Jul;34(1):69-78. PMID: 21598343.

123. Hidaka M, Takatsuki M, Okudaira S, et al. The expression of transporter OATP2/OATP8 decreases in undetectable hepatocellular carcinoma by Gd-EOB-MRI in the explanted cirrhotic liver. Hepatology International. 2012:1-7.

124. Higashihara H, Osuga K, Onishi H, et al. Diagnostic accuracy of C-arm CT during selective transcatheter angiography for hepatocellular carcinoma: comparison with intravenous contrast-enhanced,

biphasic, dynamic MDCT. Eur Radiol. 2012 Apr;22(4):872-9. PMID: 22120061.

125. Hirakawa M, Yoshimitsu K, Irie H, et al. Performance of radiological methods in diagnosing hepatocellular carcinoma preoperatively in a recipient of living related liver transplantation: comparison with step section histopathology. Jpn J Radiol. 2011 Feb;29(2):129-37. PMID: 21359938.

126. Hori M, Murakami T, Kim T, et al. Detection of hypervascular hepatocellular carcinoma: comparison of SPIO-enhanced MRI with dynamic helical CT. J Comput Assist Tomogr. 2002 Sep-Oct;26(5):701-10. PMID: 12439302.

127. Hori M, Murakami T, Oi H, et al. Sensitivity in detection of hypervascular hepatocellular carcinoma by helical CT with intra-arterial injection of contrast medium, and by helical CT and MR imaging with intravenous injection of contrast medium. Acta Radiol. 1998 Mar;39(2):144-51. PMID: 9529444.

128. Hwang J, Kim SH, Lee MW, et al. Small (<= 2 cm) hepatocellular carcinoma in patients with chronic liver disease: comparison of gadoxetic acid-enhanced 3.0 T MRI and multiphasic 64-multirow detector CT. Br J Radiol. 2012 Jul;85(1015):e314-22. PMID: 22167508.

129. Iannaccone R, Laghi A, Catalano C, et al. Hepatocellular carcinoma: role of unenhanced and delayed phase multi-detector row helical CT in patients with cirrhosis. Radiology. 2005 Feb;234(2):460-7. PMID: 15671002.

130. Ichikawa T, Kitamura T, Nakajima H, et al. Hypervascular hepatocellular carcinoma: can double arterial phase imaging with multidetector CT improve tumor depiction in the cirrhotic liver? AJR Am J Roentgenol. 2002 Sep;179(3):751-8. PMID: 12185057.

131. Ichikawa T, Saito K, Yoshioka N, et al. Detection and characterization of focal liver lesions: a Japanese phase III, multicenter comparison between gadoxetic acid disodium-enhanced magnetic resonance imaging and contrast-enhanced computed tomography predominantly in patients with hepatocellular carcinoma and chronic liver disease. Investigative Radiology. 2010 Mar;45(3):133-41. PMID: 20098330.

132. Inoue T, Kudo M, Komuta M, et al. Assessment of Gd-EOB-DTPA-enhanced MRI for HCC and dysplastic nodules and comparison of detection sensitivity versus MDCT. J Gastroenterol. 2012 Sep;47(9):1036-47. PMID: 22526270.

133. Iwazawa J, Ohue S, Hashimoto N, et al. Detection of hepatocellular carcinoma: comparison of angiographic C-arm CT and MDCT. AJR Am J Roentgenol. 2010 Oct;195(4):882-7. PMID: 20858813.

134. Jang HJ, Lim JH, Lee SJ, et al. Hepatocellular carcinoma: are combined CT during arterial portography and CT hepatic arteriography in addition to triple-phase helical CT all necessary for preoperative evaluation? Radiology. 2000 May;215(2):373-80. PMID: 10796910.

135. Jeng C-M, Kung C-H, Wang Y-C, et al. Spiral biphasic contrast-enhanced computerized tomography in the diagnosis of hepatocellular carcinoma. J Formos Med Assoc. 2002 Aug;101(8):588-92. PMID: 12440092.

136. Kang BK, Lim JH, Kim SH, et al. Preoperative depiction of hepatocellular carcinoma: ferumoxides-enhanced MR imaging versus triple-phase helical CT. Radiology. 2003 Jan;226(1):79-85. PMID: 12511672.

137. Kawaoka T, Aikata H, Takaki S, et al. FDG positron emission tomography/computed tomography for the detection of extrahepatic metastases from hepatocellular carcinoma. Hepatol Res. 2009;39(2):134-42. PMID: 19208034.

138. Kawata S, Murakami T, Kim T, et al. Multidetector CT: diagnostic impact of slice thickness on detection of hypervascular hepatocellular carcinoma. AJR Am J Roentgenol. 2002 Jul;179(1):61-6. PMID: 12076906.

139. Khan MA, Combs CS, Brunt EM, et al. Positron emission tomography scanning in the evaluation of hepatocellular carcinoma. J Hepatol. 2000 May;32(5):792-7. PMID: 10845666.

140. Kim KW, Lee JM, Klotz E, et al. Quantitative CT color mapping of the arterial enhancement fraction of the liver to detect hepatocellular carcinoma. Radiology. 2009 Feb;250(2):425-34. PMID: 19188314.

141. Kim SE, Lee HC, Shim JH, et al. Noninvasive diagnostic criteria for hepatocellular carcinoma in hepatic masses >2 cm in a hepatitis B virus-endemic area. Liver Int. 2011 Nov;31(10):1468-76. PMID: 21745284.

142. Kim SH, Choi D, Kim SH, et al. Ferucarbotran-enhanced MRI versus triple-phase MDCT for the preoperative detection of hepatocellular carcinoma. AJR Am J Roentgenol. 2005 Apr;184(4):1069-76. PMID: 15788575.

143. Kim SH, Kim SH, Lee J, et al. Gadoxetic acid-enhanced MRI versus triple-phase MDCT for the preoperative detection of hepatocellular carcinoma. AJR Am J Roentgenol. 2009 Jun;192(6):1675-81. PMID: 19457834.

144. Kim SJ, Kim SH, Lee J, et al. Ferucarbotran-enhanced 3.0-T magnetic resonance imaging using parallel imaging technique compared with triple-phase multidetector row computed tomography for the preoperative detection of hepatocellular carcinoma.[Erratum appears in J Comput Assist

Tomogr. 2008 Jul-Aug;32(4):615]. J Comput Assist Tomogr. 2008 May-Jun;32(3):379-85. PMID: 18520541.

145. Kim SK, Lim JH, Lee WJ, et al. Detection of hepatocellular carcinoma: comparison of dynamic three-phase computed tomography images and four-phase computed tomography images using multidetector row helical computed tomography. J Comput Assist Tomogr. 2002 Sep-Oct;26(5):691-8. PMID: 12439300.

146. Kim T, Murakami T, Hori M, et al. Small hypervascular hepatocellular carcinoma revealed by double arterial phase CT performed with single breath-hold scanning and automatic bolus tracking. AJR Am J Roentgenol. 2002 Apr;178(4):899-904. PMID: 11906869.

147. Kim YK, Kim CS, Chung GH, et al. Comparison of gadobenate dimeglumine-enhanced dynamic MRI and 16-MDCT for the detection of hepatocellular carcinoma. AJR Am J Roentgenol. 2006 Jan;186(1):149-57. PMID: 16357395.

148. Kim YK, Kim CS, Han YM, et al. Detection of hepatocellular carcinoma: gadoxetic acid-enhanced 3-dimensional magnetic resonance imaging versus multi-detector row computed tomography. J Comput Assist Tomogr. 2009 Nov-Dec;33(6):844-50. PMID: 19940648.

149. Kim YK, Kwak HS, Kim CS, et al. Hepatocellular carcinoma in patients with chronic liver disease: comparison of SPIO-enhanced MR imaging and 16-detector row CT. Radiology. 2006 Feb;238(2):531-41. PMID: 16371577.

150. Kitamura T, Ichikawa T, Erturk SM, et al. Detection of hypervascular hepatocellular carcinoma with multidetector-row CT: single arterial-phase imaging with computer-assisted automatic bolus-tracking technique compared with double arterial-phase imaging. J Comput Assist Tomogr. 2008 Sep-Oct;32(5):724-9. PMID: 18830101.

151. Kumano S, Uemura M, Haraikawa T, et al. Efficacy of double arterial phase dynamic magnetic resonance imaging with the sensitivity encoding technique versus dynamic multidetector-row helical computed tomography for detecting hypervascular hepatocellular carcinoma. Jpn J Radiol. 2009 Jul;27(6):229-36. PMID: 19626408.

152. Laghi A, Iannaccone R, Rossi P, et al. Hepatocellular carcinoma: detection with triple-phase multi-detector row helical CT in patients with chronic hepatitis. Radiology. 2003 Feb;226(2):543-9. PMID: 12563152.

153. Lee CH, Kim KA, Lee J, et al. Using low tube voltage (80kVp) quadruple phase liver CT for the detection of hepatocellular carcinoma: two-year experience and comparison with Gd-EOB-DTPA enhanced liver MRI. Eur J Radiol. 2012 Apr;81(4):e605-11. PMID: 22297180.

154. Lee DH, Kim SH, Lee JM, et al. Diagnostic performance of multidetector row computed tomography, superparamagnetic iron oxide-enhanced magnetic resonance imaging, and dual-contrast magnetic resonance imaging in predicting the appropriateness of a transplant recipient based on milan criteria: correlation with histopathological findings. Investigative Radiology. 2009 Jun;44(6):311-21. PMID: 19462486.

155. Lee J, Won JL, Hyo KL, et al. Early hepatocellular carcinoma: Three-phase helical CT features of 16 patients. Korean J Radiol. 2008;9(4):325-32. PMID: 18682670.

156. Lee JE, Jang JY, Jeong SW, et al. Diagnostic value for extrahepatic metastases of hepatocellular carcinoma in positron emission tomography/computed tomography scan. World J Gastroenterol. 2012 Jun 21;18(23):2979-87. PMID: 22736922.

157. Lee J-M, Kim I-H, Kwak H-S, et al. Detection of small hypervascular hepatocellular carcinomas in cirrhotic patients: comparison of superparamagnetic iron oxide-enhanced MR imaging with dual-phase spiral CT. Korean J Radiol. 2003 Jan-Mar;4(1):1-8. PMID: 12679628.

158. Li CS, Chen RC, Tu HY, et al. Imaging well-differentiated hepatocellular carcinoma with dynamic triple-phase helical computed tomography. Br J Radiol. 2006 Aug;79(944):659-65. PMID: 16641423.

159. Lim JH, Choi D, Kim SH, et al. Detection of hepatocellular carcinoma: value of adding delayed phase imaging to dual-phase helical CT. AJR Am J Roentgenol. 2002 Jul;179(1):67-73. PMID: 12076907.

160. Lim JH, Kim CK, Lee WJ, et al. Detection of hepatocellular carcinomas and dysplastic nodules in cirrhotic livers: accuracy of helical CT in transplant patients. AJR Am J Roentgenol. 2000 Sep;175(3):693-8. PMID: 10954452.

161. Lin M-T, Chen C-L, Wang C-C, et al. Diagnostic sensitivity of hepatocellular carcinoma imaging and its application to non-cirrhotic patients. J Gastroenterol Hepatol. 2011 Apr;26(4):745-50. PMID: 21418303.

162. Liu YI, Kamaya A, Jeffrey RB, et al. Multidetector computed tomography triphasic evaluation of the liver before transplantation: importance of equilibrium phase washout and morphology for characterizing hypervascular lesions. J Comput Assist Tomogr. 2012 Mar-Apr;36(2):213-9. PMID: 22446362.

163. Liu YI, Shin LK, Jeffrey RB, et al. Quantitatively defining washout in hepatocellular carcinoma. AJR Am J Roentgenol. 2013 Jan;200(1):84-9. PMID: 23255745.

164. Lu CH, Chen CL, Cheng YF, et al. Correlation between imaging and pathologic findings in explanted livers of hepatocellular carcinoma cases. Transplant Proc. 2010 Apr;42(3):830-3. PMID: 20430183.

165. Lv P, Lin XZ, Chen K, et al. Spectral CT in patients with small HCC: investigation of image quality and diagnostic accuracy. Eur Radiol. 2012 Oct;22(10):2117-24. PMID: 22618521.

166. Lv P, Lin XZ, Li J, et al. Differentiation of small hepatic hemangioma from small hepatocellular carcinoma: recently introduced spectral CT method. Radiology. 2011 Jun;259(3):720-9. PMID: 21357524.

167. Maetani YS, Ueda M, Haga H, et al. Hepatocellular carcinoma in patients undergoing living-donor liver transplantation. Accuracy of multidetector computed tomography by viewing images on digital monitors. Intervirology. 2008;51 Suppl 1:46-51. PMID: 18544948.

168. Marin D, Catalano C, De Filippis G, et al. Detection of hepatocellular carcinoma in patients with cirrhosis: added value of coronal reformations from isotropic voxels with 64-MDCT. AJR Am J Roentgenol. 2009 Jan;192(1):180-7. PMID: 19098199.

169. Marin D, Di Martino M, Guerrisi A, et al. Hepatocellular carcinoma in patients with cirrhosis: qualitative comparison of gadobenate dimeglumine-enhanced MR imaging and multiphasic 64-section CT. Radiology. 2009 Apr;251(1):85-95. PMID: 19332848.

170. Monzawa S, Ichikawa T, Nakajima H, et al. Dynamic CT for detecting small hepatocellular carcinoma: usefulness of delayed phase imaging. AJR Am J Roentgenol. 2007 Jan;188(1):147-53. PMID: 17179357.

171. Mortele KJ, De Keukeleire K, Praet M, et al. Malignant focal hepatic lesions complicating underlying liver disease: dual-phase contrast-enhanced spiral CT sensitivity and specificity in orthotopic liver transplant patients. Eur Radiol. 2001;11(9):1631-8. PMID: 11511882.

172. Murakami T, Kim T, Kawata S, et al. Evaluation of optimal timing of arterial phase imaging for the detection of hypervascular hepatocellular carcinoma by using triple arterial phase imaging with multidetector-row helical computed tomography. Investigative Radiology. 2003 Aug;38(8):497-503. PMID: 12874516.

173. Murakami T, Kim T, Takamura M, et al. Hypervascular hepatocellular carcinoma: detection with double arterial phase multi-detector row helical CT. Radiology. 2001 Mar;218(3):763-7. PMID: 11230652.

174. Nagaoka S, Itano S, Ishibashi M, et al. Value of fusing PET plus CT images in hepatocellular carcinoma and combined hepatocellular and cholangiocarcinoma patients with extrahepatic metastases: preliminary findings. Liver Int. 2006 Sep;26(7):781-8. PMID: 16911459.

175. Nakamura H, Ito N, Kotake F, et al. Tumor-detecting capacity and clinical usefulness of SPIO-MRI in patients with hepatocellular carcinoma. J Gastroenterol. 2000;35(11):849-55. PMID: 11085494.

176. Nakamura Y, Tashiro H, Nambu J, et al. Detectability of hepatocellular carcinoma by gadoxetate disodium-enhanced hepatic MRI: Tumor-by-tumor analysis in explant livers. J Magn Reson Imaging. 2013;37(3):684-91. PMID: 23055436.

177. Nakayama A, Imamura H, Matsuyama Y, et al. Value of lipiodol computed tomography and digital subtraction angiography in the era of helical biphasic computed tomography as preoperative assessment of hepatocellular carcinoma. Ann Surg. 2001 Jul;234(1):56-62. PMID: 11420483.

178. Noguchi Y, Murakami T, Kim T, et al. Detection of hepatocellular carcinoma: comparison of dynamic MR imaging with dynamic double arterial phase helical CT. AJR Am J Roentgenol. 2003 Feb;180(2):455-60. PMID: 12540451.

179. Noguchi Y, Murakami T, Kim T, et al. Detection of hypervascular hepatocellular carcinoma by dynamic magnetic resonance imaging with double-echo chemical shift in-phase and opposed-phase gradient echo technique: comparison with dynamic helical computed tomography imaging with double arterial phase. J Comput Assist Tomogr. 2002 Nov-Dec;26(6):981-7. PMID: 12488747.

180. Onishi H, Kim T, Imai Y, et al. Hypervascular hepatocellular carcinomas: detection with gadoxetate disodium-enhanced MR imaging and multiphasic multidetector CT. Eur Radiol. 2012 Apr;22(4):845-54. PMID: 22057248.

181. Park JH, Kim SH, Park HS, et al. Added value of 80 kVp images to averaged 120 kVp images in the detection of hepatocellular carcinomas in liver transplantation candidates using dual-source dual-energy MDCT: results of JAFROC analysis. Eur J Radiol. 2011 Nov;80(2):e76-85. PMID: 20875937.

182. Peterson MS, Baron RL, Marsh JW, Jr., et al. Pretransplantation surveillance for possible hepatocellular carcinoma in patients with cirrhosis: epidemiology and CT-based tumor detection rate in 430 cases with surgical pathologic correlation. Radiology. 2000 Dec;217(3):743-9. PMID: 11110938.

183. Pitton MB, Kloeckner R, Herber S, et al. MRI versus 64-row MDCT for diagnosis of hepatocellular carcinoma. World J Gastroenterol. 2009 Dec 28;15(48):6044-51. PMID: 20027676.

184. Pozzi Mucelli RM, Como G, Del Frate C, et al. Multidetector CT with double arterial phase and high-iodine-concentration contrast agent in the detection of hepatocellular carcinoma. Radiol Med (Torino). 2006 Mar;111(2):181-91. PMID: 16671376.

185. Pugacheva O, Matsui O, Kozaka K, et al. Detection of small hypervascular hepatocellular carcinomas by EASL criteria: comparison with double-phase CT during hepatic arteriography. Eur J Radiol. 2011 Dec;80(3):e201-6. PMID: 20855175.

186. Rizvi S, Camci C, Yong Y, et al. Is post-Lipiodol CT better than i.v. contrast CT scan for early detection of HCC? A single liver transplant center experience. Transplant Proc. 2006 Nov;38(9):2993-5. PMID: 17112883.

187. Ronzoni A, Artioli D, Scardina R, et al. Role of MDCT in the diagnosis of hepatocellular carcinoma in patients with cirrhosis undergoing orthotopic liver transplantation. AJR Am J Roentgenol. 2007 Oct;189(4):792-8. PMID: 17885047.

188. Sano K, Ichikawa T, Motosugi U, et al. Imaging study of early hepatocellular carcinoma: usefulness of gadoxetic acid-enhanced MR imaging. Radiology. 2011 Dec;261(3):834-44. PMID: 21998047.

189. Schima W, Hammerstingl R, Catalano C, et al. Quadruple-phase MDCT of the liver in patients with suspected hepatocellular carcinoma: Effect of contrast material flow rate. AJR Am J Roentgenol. 2006;186(6):1571-9. PMID: 16714645.

190. Sofue K, Tsurusaki M, Kawasaki R, et al. Evaluation of hypervascular hepatocellular carcinoma in cirrhotic liver: comparison of different concentrations of contrast material with multi-detector row helical CT--a prospective randomized study. Eur J Radiol. 2011 Dec;80(3):e237-42. PMID: 21067880.

191. Sun HY, Lee JM, Shin CI, et al. Gadoxetic acid-enhanced magnetic resonance imaging for differentiating small hepatocellular carcinomas (< or =2 cm in diameter) from arterial enhancing pseudolesions: special emphasis on hepatobiliary phase imaging. Investigative Radiology. 2010 Feb;45(2):96-103. PMID: 20057319.

192. Valls C, Cos M, Figueras J, et al. Pretransplantation diagnosis and staging of hepatocellular carcinoma in patients with cirrhosis: value of dual-phase helical CT. AJR Am J Roentgenol. 2004 Apr;182(4):1011-7. PMID: 15039179.

193. Wagnetz U, Atri M, Massey C, et al. Intraoperative ultrasound of the liver in primary and secondary hepatic malignancies: comparison with preoperative 1.5-T MRI and 64-MDCT. AJR Am J Roentgenol. 2011 Mar;196(3):562-8. PMID: 21343497.

194. Xiao X-g, Han X, Shan W-d, et al. Multi-slice CT angiography by triple-phase enhancement in preoperative evaluation of hepatocellular carcinoma. Chinese Medical Journal. 2005 May 20;118(10):844-9. PMID: 15989766.

195. Yan FH, Shen JZ, Li RC, et al. Enhancement patterns of small hepatocellular carcinoma shown by dynamic MRI and CT. Hepatobiliary Pancreat Dis Int. 2002 Aug;1(3):420-4. PMID: 14607719.

196. Yu Y, Lin X, Chen K, et al. Hepatocellular carcinoma and focal nodular hyperplasia of the liver: differentiation with CT spectral imaging. Eur Radiol. 2013 Jun;23(6):1660-8 Epub 2013 Jan 10. PMID: 23306709.

197. Yukisawa S, Okugawa H, Masuya Y, et al. Multidetector helical CT plus superparamagnetic iron oxide-enhanced MR imaging for focal hepatic lesions in cirrhotic liver: a comparison with multi-phase CT during hepatic arteriography. Eur J Radiol. 2007 Feb;61(2):279-89. PMID: 17070663.

198. Zacherl J, Pokieser P, Wrba F, et al. Accuracy of multiphasic helical computed tomography and intraoperative sonography in patients undergoing orthotopic liver transplantation for hepatoma: what is the truth? Ann Surg. 2002 Apr;235(4):528-32. PMID: 11923609.

199. Zhao H, Yao JL, Han MJ, et al. Multiphase hepatic scans with multirow-detector helical CT in detection of hypervascular hepatocellular carcinoma. Hepatobiliary Pancreat Dis Int. 2004 May;3(2):204-8. PMID: 15138110.

200. Zhao H, Yao J-L, Wang Y, et al. Detection of small hepatocellular carcinoma: comparison of dynamic enhancement magnetic resonance imaging and multiphase multirow-detector helical CT scanning. World J Gastroenterol. 2007 Feb 28;13(8):1252-6. PMID: 17451209.

201. Zhao H, Zhou K-R, Yan F-H. Role of multiphase scans by multirow-detector helical CT in detecting small hepatocellular carcinoma. World J Gastroenterol. 2003 Oct;9(10):2198-201. PMID: 14562377.

202. Zheng X-H, Guan Y-S, Zhou X-P, et al. Detection of hypervascular hepatocellular carcinoma: Comparison of multi-detector CT with digital subtraction angiography and Lipiodol CT. World J Gastroenterol. 2005 Jan 14;11(2):200-3. PMID: 15633215.

203. Ahn SS, Kim M-J, Lim JS, et al. Added value of gadoxetic acid-enhanced hepatobiliary phase MR imaging in the diagnosis of hepatocellular carcinoma. Radiology. 2010 May;255(2):459-66. PMID: 20413759.

204. An C, Park MS, Kim D, et al. Added value of subtraction imaging in detecting arterial enhancement in small (<3 cm) hepatic nodules on dynamic contrast-enhanced MRI in patients at high

risk of hepatocellular carcinoma. Eur Radiol. 2013;23:924-30. PMID: 23138382.

205. An C, Park M-S, Jeon H-M, et al. Prediction of the histopathological grade of hepatocellular carcinoma using qualitative diffusion-weighted, dynamic, and hepatobiliary phase MRI. Eur Radiol. 2012 Aug;22(8):1701-8. PMID: 22434421.

206. Cereser L, Furlan A, Bagatto D, et al. Comparison of portal venous and delayed phases of gadolinium-enhanced magnetic resonance imaging study of cirrhotic liver for the detection of contrast washout of hypervascular hepatocellular carcinoma. J Comput Assist Tomogr. 2010;34(5):706-11.

207. Choi D, Kim SH, Lim JH, et al. Detection of hepatocellular carcinoma: combined T2-weighted and dynamic gadolinium-enhanced MRI versus combined CT during arterial portography and CT hepatic arteriography. J Comput Assist Tomogr. 2001 Sep-Oct;25(5):777-85. PMID: 11584240.

208. Choi SH, Lee JM, Yu NC, et al. Hepatocellular carcinoma in liver transplantation candidates: detection with gadobenate dimeglumine-enhanced MRI. AJR Am J Roentgenol. 2008 Aug;191(2):529-36. PMID: 18647927.

209. Chou C-T, Chen R-C, Chen W-T, et al. Characterization of hyperintense nodules on T1-weighted liver magnetic resonance imaging: comparison of Ferucarbotran-enhanced MRI with accumulation-phase FS-T1WI and gadolinium-enhanced MRI. J Chin Med Assoc. 2011 Feb;74(2):62-8. PMID: 21354082.

210. Chung J, Yu J-S, Kim DJ, et al. Hypervascular hepatocellular carcinoma in the cirrhotic liver: diffusion-weighted imaging versus superparamagnetic iron oxide-enhanced MRI. Magn Reson Imaging. 2011 Nov;29(9):1235-43. PMID: 21907517.

211. Chung S-H, Kim M-J, Choi J-Y, et al. Comparison of two different injection rates of gadoxetic acid for arterial phase MRI of the liver. J Magn Reson Imaging. 2010 Feb;31(2):365-72. PMID: 20099350.

212. Goshima S, Kanematsu M, Matsuo M, et al. Nodule-in-nodule appearance of hepatocellular carcinomas: comparison of gadolinium-enhanced and ferumoxides-enhanced magnetic resonance imaging. J Magn Reson Imaging. 2004 Aug;20(2):250-5. PMID: 15269950.

213. Guo L, Liang C, Yu T, et al. 3 T MRI of hepatocellular carcinomas in patients with cirrhosis: does T2-weighted imaging provide added value? Clin Radiol. 2012 Apr;67(4):319-28. PMID: 22099524.

214. Hardie AD, Kizziah MK, Boulter DJ. Diagnostic accuracy of diffusion-weighted MRI for identifying hepatocellular carcinoma with liver explant correlation. J Med Imaging Radiat Oncol. 2011 Aug;55(4):362-7. PMID: 21843170.

215. Hardie AD, Kizziah MK, Rissing MS. Can the patient with cirrhosis be imaged for hepatocellular carcinoma without gadolinium?: Comparison of combined T2-weighted, T2*-weighted, and diffusion-weighted MRI with gadolinium-enhanced MRI using liver explantation standard. J Comput Assist Tomogr. 2011 Nov-Dec;35(6):711-5. PMID: 22082541.

216. Hardie AD, Nance JW, Boulter DJ, et al. Assessment of the diagnostic accuracy of T2*-weighted MR imaging for identifying hepatocellular carcinoma with liver explant correlation. Eur J Radiol. 2011 Dec;80(3):e249-52. PMID: 21112710.

217. Hecht EM, Holland AE, Israel GM, et al. Hepatocellular carcinoma in the cirrhotic liver: gadolinium-enhanced 3D T1-weighted MR imaging as a stand-alone sequence for diagnosis. Radiology. 2006 May;239(2):438-47. PMID: 16641353.

218. Ito K, Fujita T, Shimizu A, et al. Multiarterial phase dynamic MRI of small early enhancing hepatic lesions in cirrhosis or chronic hepatitis: Differentiating between hypervascular hepatocellular carcinomas and pseudolesions. AJR Am J Roentgenol. 2004;183(3):699-705. PMID: 15333358.

219. Jeong MG, Yu JS, Kim KW, et al. Early homogeneously enhancing hemangioma versus hepatocellular carcinoma: differentiation using quantitative analysis of multiphasic dynamic magnetic resonance imaging. Yonsei Med J. 1999 Jun;40(3):248-55. PMID: 10412337.

220. Jeong WK, Byun JH, Lee SS, et al. Gadobenate dimeglumine-enhanced liver MR imaging in cirrhotic patients: quantitative and qualitative comparison of 1-hour and 3-hour delayed images. J Magn Reson Imaging. 2011 Apr;33(4):889-97. PMID: 21448954.

221. Jin Y, Nah S, Lee J, et al. Utility of Adding Primovist Magnetic Resonance Imaging to Analysis of Hepatocellular Carcinoma by Liver Dynamic Computed Tomography. Clin Gastroenterol Hepatol. 2013;11(2):187-92. PMID: 23142203.

222. Kamura T, Kimura M, Sakai K, et al. Small hypervascular hepatocellular carcinoma versus hypervascular pseudolesions: Differential diagnosis on MRI. Abdom Imaging. 2002;27(3):315-24. PMID: 12173363.

223. Kim AY, Kim YK, Lee MW, et al. Detection of hepatocellular carcinoma in gadoxetic acid-enhanced MRI and diffusion-weighted MRI with respect to the severity of liver cirrhosis. Acta Radiol. 2012 Oct 1;53(8):830-8. PMID: 22847903.

224. Kim DJ, Yu JS, Kim JH, et al. Small hypervascular hepatocellular carcinomas: value of diffusion-

weighted imaging compared with "washout" appearance on dynamic MRI. Br J Radiol. 2012 Oct;85(1018):e879-86. PMID: 22573299.

225. Kim M-J, Lee M, Choi J-Y, et al. Imaging features of small hepatocellular carcinomas with microvascular invasion on gadoxetic acid-enhanced MR imaging. Eur J Radiol. 2012 Oct;81(10):2507-12. PMID: 22137613.

226. Kim TK, Lee KH, Jang H-J, et al. Analysis of gadobenate dimeglumine-enhanced MR findings for characterizing small (1-2-cm) hepatic nodules in patients at high risk for hepatocellular carcinoma. Radiology. 2011 Jun;259(3):730-8. PMID: 21364083.

227. Kim YK, Kim CS, Han YM, et al. Detection of liver malignancy with gadoxetic acid-enhanced MRI: is addition of diffusion-weighted MRI beneficial?.[Erratum appears in Clin Radiol. 2011 Oct;66(10):1006]. Clin Radiol. 2011 Jun;66(6):489-96. PMID: 21367403.

228. Kim YK, Kim CS, Han YM, et al. Detection of small hepatocellular carcinoma: can gadoxetic acid-enhanced magnetic resonance imaging replace combining gadopentetate dimeglumine-enhanced and superparamagnetic iron oxide-enhanced magnetic resonance imaging? Investigative Radiology. 2010 Nov;45(11):740-6. PMID: 20644488.

229. Kim YK, Kim CS, Han YM, et al. Comparison of gadoxetic acid-enhanced MRI and superparamagnetic iron oxide-enhanced MRI for the detection of hepatocellular carcinoma. Clin Radiol. 2010 May;65(5):358-65. PMID: 20380933.

230. Kim YK, Kim CS, Han YM, et al. Detection of small hepatocellular carcinoma: intraindividual comparison of gadoxetic acid-enhanced MRI at 3.0 and 1.5 T.[Erratum appears in Invest Radiol. 2011 Sep;46(9):600]. Investigative Radiology. 2011 Jun;46(6):383-9. PMID: 21467946.

231. Kim YK, Kim CS, Kwak HS, et al. Three-dimensional dynamic liver MR imaging using sensitivity encoding for detection of hepatocellular carcinomas: comparison with superparamagnetic iron oxide-enhanced mr imaging. J Magn Reson Imaging. 2004 Nov;20(5):826-37. PMID: 15503325.

232. Kim YK, Kwak HS, Han YM, et al. Usefulness of combining sequentially acquired gadobenate dimeglumine-enhanced magnetic resonance imaging and resovist-enhanced magnetic resonance imaging for the detection of hepatocellular carcinoma: comparison with computed tomography hepatic arteriography and computed tomography arterioportography using 16-slice multidetector computed tomography. J Comput Assist Tomogr. 2007 Sep-Oct;31(5):702-11. PMID: 17895780.

233. Kim YK, Lee YH, Kim CS, et al. Added diagnostic value of T2-weighted MR imaging to gadolinium-enhanced three-dimensional dynamic MR imaging for the detection of small hepatocellular carcinomas. Eur J Radiol. 2008 Aug;67(2):304-10. PMID: 17714904.

234. Kim YK, Lee YH, Kim CS, et al. Double-dose 1.0-M gadobutrol versus standard-dose 0.5-M gadopentetate dimeglumine in revealing small hypervascular hepatocellular carcinomas. Eur Radiol. 2008 Jan;18(1):70-7. PMID: 17404740.

235. Kondo H, Kanematsu M, Itoh K, et al. Does T2-weighted MR imaging improve preoperative detection of malignant hepatic tumors? Observer performance study in 49 surgically proven cases. Magn Reson Imaging. 2005 Jan;23(1):89-95. PMID: 15733793.

236. Koushima Y, Ebara M, Fukuda H, et al. Small hepatocellular carcinoma: assessment with T1-weighted spin-echo magnetic resonance imaging with and without fat suppression. Eur J Radiol. 2002 Jan;41(1):34-41. PMID: 11750150.

237. Krinsky GA, Lee VS, Theise ND, et al. Transplantation for hepatocellular carcinoma and cirrhosis: Sensitivity of magnetic resonance imaging. Liver Transpl. 2002;8(12):1156-64. PMID: 12474156.

238. Krinsky GA, Lee VS, Theise ND, et al. Hepatocellular carcinoma and dysplastic nodules in patients with cirrhosis: prospective diagnosis with MR imaging and explantation correlation. Radiology. 2001 May;219(2):445-54. PMID: 11323471.

239. Kwak H-S, Lee J-M, Kim C-S. Preoperative detection of hepatocellular carcinoma: comparison of combined contrast-enhanced MR imaging and combined CT during arterial portography and CT hepatic arteriography. Eur Radiol. 2004 Mar;14(3):447-57. PMID: 14531005.

240. Kwak H-S, Lee J-M, Kim Y-K, et al. Detection of hepatocellular carcinoma: comparison of ferumoxides-enhanced and gadolinium-enhanced dynamic three-dimensional volume interpolated breath-hold MR imaging. Eur Radiol. 2005 Jan;15(1):140-7. PMID: 15449000.

241. Lauenstein TC, Salman K, Morreira R, et al. Gadolinium-enhanced MRI for tumor surveillance before liver transplantation: center-based experience. AJR Am J Roentgenol. 2007 Sep;189(3):663-70. PMID: 17715115.

242. Lee JY, Kim SH, Jeon YH, et al. Ferucarbotran-enhanced magnetic resonance imaging versus gadoxetic acid-enhanced magnetic resonance imaging for the preoperative detection of hepatocellular carcinoma: initial experience. J Comput Assist Tomogr. 2010 Jan;34(1):127-34. PMID: 20118735.

243. Lee MH, Kim SH, Park MJ, et al. Gadoxetic acid-enhanced hepatobiliary phase MRI and high-b-value diffusion-weighted imaging to distinguish well-differentiated hepatocellular carcinomas from benign nodules in patients with chronic liver disease. AJR Am J Roentgenol. 2011 Nov;197(5):W868-75. PMID: 22021534.

244. Marrero JA, Hussain HK, Nghiem HV, et al. Improving the prediction of hepatocellular carcinoma in cirrhotic patients with an arterially-enhancing liver mass. Liver Transpl. 2005 Mar;11(3):281-9. PMID: 15719410.

245. Matsuo M, Kanematsu M, Itoh K, et al. Detection of malignant hepatic tumors: comparison of gadolinium-and ferumoxide-enhanced MR imaging. AJR Am J Roentgenol. 2001 Sep;177(3):637-43. PMID: 11517061.

246. Mori K, Yoshioka H, Takahashi N, et al. Triple arterial phase dynamic MRI with sensitivity encoding for hypervascular hepatocellular carcinoma: comparison of the diagnostic accuracy among the early, middle, late, and whole triple arterial phase imaging. AJR Am J Roentgenol. 2005 Jan;184(1):63-9. PMID: 15615952.

247. Motosugi U, Ichikawa T, Sou H, et al. Distinguishing hypervascular pseudolesions of the liver from hypervascular hepatocellular carcinomas with gadoxetic acid-enhanced MR imaging. Radiology. 2010 Jul;256(1):151-8. PMID: 20574092.

248. Ooka Y, Kanai F, Okabe S, et al. Gadoxetic acid-enhanced MRI compared with CT during angiography in the diagnosis of hepatocellular carcinoma. Magn Reson Imaging. 2013;31(5):748-54 Epub 2012 Dec 5. PMID: 23218794.

249. Park G, Kim YK, Kim CS, et al. Diagnostic efficacy of gadoxetic acid-enhanced MRI in the detection of hepatocellular carcinomas: comparison with gadopentetate dimeglumine. Br J Radiol. 2010 Dec;83(996):1010-6. PMID: 20682591.

250. Park M-S, Kim S, Patel J, et al. Hepatocellular carcinoma: detection with diffusion-weighted versus contrast-enhanced magnetic resonance imaging in pretransplant patients. Hepatology. 2012 Jul;56(1):140-8. PMID: 22370974.

251. Park Y, Kim SH, Kim SH, et al. Gadoxetic acid (Gd-EOB-DTPA)-enhanced MRI versus gadobenate dimeglumine (Gd-BOPTA)-enhanced MRI for preoperatively detecting hepatocellular carcinoma: an initial experience. Korean J Radiol. 2010 Jul-Aug;11(4):433-40. PMID: 20592927.

252. Pauleit D, Textor J, Bachmann R, et al. Hepatocellular carcinoma: detection with gadolinium- and ferumoxides-enhanced MR imaging of the liver. Radiology. 2002 Jan;222(1):73-80. PMID: 11756708.

253. Petruzzi N, Mitchell D, Guglielmo F, et al. Hepatocellular carcinoma likelihood on MRI exams. Evaluation of a standardized categorization System. Acad Radiol. 2013 Jun;20(6):694-8 Epub 2013 Mar 28. PMID: 23541479.

254. Piana G, Trinquart L, Meskine N, et al. New MR imaging criteria with a diffusion-weighted sequence for the diagnosis of hepatocellular carcinoma in chronic liver diseases. J Hepatol. 2011 Jul;55(1):126-32. PMID: 21145857.

255. Rhee H, Kim MJ, Park MS, et al. Differentiation of early hepatocellular carcinoma from benign hepatocellular nodules on gadoxetic acid-enhanced MRI. Br J Radiol. 2012 Oct;85(1018):e837-44. PMID: 22553295.

256. Rimola J, Forner A, Tremosini S, et al. Non-invasive diagnosis of hepatocellular carcinoma <= 2 cm in cirrhosis. Diagnostic accuracy assessing fat, capsule and signal intensity at dynamic MRI. J Hepatol. 2012 Jun;56(6):1317-23. PMID: 22314420.

257. Secil M, Obuz F, Altay C, et al. The role of dynamic subtraction MRI in detection of hepatocellular carcinoma. Diagn Interv Radiol. 2008 Dec;14(4):200-4. PMID: 19061165.

258. Simon G, Link TM, Wortler K, et al. Detection of hepatocellular carcinoma: comparison of Gd-DTPA- and ferumoxides-enhanced MR imaging. Eur Radiol. 2005 May;15(5):895-903. PMID: 15800773.

259. Suh YJ, Kim M-J, Choi J-Y, et al. Differentiation of hepatic hyperintense lesions seen on gadoxetic acid-enhanced hepatobiliary phase MRI. AJR Am J Roentgenol. 2011 Jul;197(1):W44-52. PMID: 21700994.

260. Tanaka O, Ito H, Yamada K, et al. Higher lesion conspicuity for SENSE dynamic MRI in detecting hypervascular hepatocellular carcinoma: analysis through the measurements of liver SNR and lesion-liver CNR comparison with conventional dynamic MRI. Eur Radiol. 2005 Dec;15(12):2427-34. PMID: 16041592.

261. Tang Y, Yamashita Y, Arakawa A, et al. Detection of hepatocellular carcinoma arising in cirrhotic livers: comparison of gadolinium- and ferumoxides-enhanced MR imaging. AJR Am J Roentgenol. 1999 Jun;172(6):1547-54. PMID: 10350287.

262. Tanimoto A, Yuasa Y, Jinzaki M, et al. Routine MR imaging protocol with breath-hold fast scans: diagnostic efficacy for focal liver lesions. Radiat Med. 2002 Jul-Aug;20(4):169-79. PMID: 12296432.

263. Tsurusaki M, Semelka RC, Uotani K, et al. Prospective comparison of high- and low-spatial-resolution dynamic MR imaging with sensitivity encoding (SENSE) for hypervascular hepatocellular

carcinoma. Eur Radiol. 2008 Oct;18(10):2206-12. PMID: 18446347.

264. Vandecaveye V, De Keyzer F, Verslype C, et al. Diffusion-weighted MRI provides additional value to conventional dynamic contrast-enhanced MRI for detection of hepatocellular carcinoma. Eur Radiol. 2009 Oct;19(10):2456-66. PMID: 19440718.

265. Xu P-J, Yan F-H, Wang J-H, et al. Added value of breathhold diffusion-weighted MRI in detection of small hepatocellular carcinoma lesions compared with dynamic contrast-enhanced MRI alone using receiver operating characteristic curve analysis. J Magn Reson Imaging. 2009 Feb;29(2):341-9. PMID: 19161186.

266. Xu P-J, Yan F-H, Wang J-H, et al. Contribution of diffusion-weighted magnetic resonance imaging in the characterization of hepatocellular carcinomas and dysplastic nodules in cirrhotic liver. J Comput Assist Tomogr. 2010 Jul;34(4):506-12. PMID: 20657216.

267. Yoshioka H, Takahashi N, Yamaguchi M, et al. Double arterial phase dynamic MRI with sensitivity encoding (SENSE) for hypervascular hepatocellular carcinomas. J Magn Reson Imaging. 2002 Sep;16(3):259-66. PMID: 12205581.

268. Yu JS, Chung JJ, Kim JH, et al. Small hypervascular hepatocellular carcinomas: value of "washout" on gadolinium-enhanced dynamic MR imaging compared to superparamagnetic iron oxide-enhanced imaging. Eur Radiol. 2009 Nov;19(11):2614-22. PMID: 19513719.

269. Yu JS, Kim KW, Park MS, et al. Transient peritumoral enhancement during dynamic MRI of the liver: Cavernous hemangioma versus hepatocellular carcinoma. J Comput Assist Tomogr. 2002 May-Jun;26(3):411-7. PMID: 12016371.

270. Yu JS, Lee JH, Chung JJ, et al. Small hypervascular hepatocellular carcinoma: limited value of portal and delayed phases on dynamic magnetic resonance imaging. Acta Radiol. 2008 Sep;49(7):735-43. PMID: 18608015.

271. Chen Y-K, Hsieh D-S, Liao C-S, et al. Utility of FDG-PET for investigating unexplained serum AFP elevation in patients with suspected hepatocellular carcinoma recurrence. Anticancer Res. 2005 Nov-Dec;25(6C):4719-25. PMID: 16334166.

272. Cheung TT, Chan SC, Ho CL, et al. Can positron emission tomography with the dual tracers [11 C]acetate and [18 F]fludeoxyglucose predict microvascular invasion in hepatocellular carcinoma? Liver Transpl. 2011 Oct;17(10):1218-25. PMID: 21688383.

273. Delbeke D, Martin WH, Sandler MP, et al. Evaluation of benign vs malignant hepatic lesions with positron emission tomography. Arch Surg. 1998 May;133(5):510-5; discussion 5-6. PMID: 9605913.

274. Eckel F, Herrmann K, Schmidt S, et al. Imaging of proliferation in hepatocellular carcinoma with the in vivo marker 18F-fluorothymidine. J Nucl Med. 2009 Sep;50(9):1441-7. PMID: 19690030.

275. Ho C-l, Chen S, Yeung DWC, et al. Dual-tracer PET/CT imaging in evaluation of metastatic hepatocellular carcinoma. J Nucl Med. 2007 Jun;48(6):902-9. PMID: 17504862.

276. Ho C-L, Yu SCH, Yeung DWC. 11C-acetate PET imaging in hepatocellular carcinoma and other liver masses. J Nucl Med. 2003 Feb;44(2):213-21. PMID: 12571212.

277. Hwang KH, Choi D-J, Lee S-Y, et al. Evaluation of patients with hepatocellular carcinomas using [(11)C]acetate and [(18)F]FDG PET/CT: A preliminary study. Appl Radiat Isot. 2009 Jul-Aug;67(7-8):1195-8. PMID: 19342249.

278. Jeng L-B, Changlai S-P, Shen Y-Y, et al. Limited value of 18F-2-deoxyglucose positron emission tomography to detect hepatocellular carcinoma in hepatitis B virus carriers. Hepatogastroenterology. 2003 Nov-Dec;50(54):2154-6. PMID: 14696485.

279. Kim Y-K, Lee K-W, Cho SY, et al. Usefulness 18F-FDG positron emission tomography/computed tomography for detecting recurrence of hepatocellular carcinoma in posttransplant patients. Liver Transpl. 2010 Jun;16(6):767-72. PMID: 20517911.

280. Li S, Beheshti M, Peck-Radosavljevic M, et al. Comparison of (11)C-acetate positron emission tomography and (67)Gallium citrate scintigraphy in patients with hepatocellular carcinoma. Liver Int. 2006 Oct;26(8):920-7. PMID: 16953831.

281. Liangpunsakul S, Agarwal D, Horlander JC, et al. Positron emission tomography for detecting occult hepatocellular carcinoma in hepatitis C cirrhotics awaiting for liver transplantation. Transplant Proc. 2003 Dec;35(8):2995-7. PMID: 14697959.

282. Lin W-Y, Tsai S-C, Hung G-U. Value of delayed 18F-FDG-PET imaging in the detection of hepatocellular carcinoma. Nuclear Medicine Communications. 2005 Apr;26(4):315-21. PMID: 15753790.

283. Park J-W, Kim JH, Kim SK, et al. A prospective evaluation of 18F-FDG and 11C-acetate PET/CT for detection of primary and metastatic hepatocellular carcinoma. J Nucl Med. 2008 Dec;49(12):1912-21. PMID: 18997056.

284. Sørensen M, Frisch K, Bender D, et al. The potential use of 2-[18F]fluoro-2-deoxy-D-galactose as a PET/CT tracer for detection of hepatocellular carcinoma. Eur J Nucl Med Mol Imaging. 2011;38(9):1723-31. PMID: 21553087.

285. Sugiyama M, Sakahara H, Torizuka T, et al. 18F-FDG PET in the detection of extrahepatic

metastases from hepatocellular carcinoma. J Gastroenterol. 2004;39(10):961-8.

286. Sun L, Guan YS, Pan WM, et al. Metabolic restaging of hepatocellular carcinoma using whole-body F-FDG PET/CT. World J Hepatol. 2009 Oct 31;1(1):90-7. PMID: 21160970.

287. Talbot J, Fartoux L, Balogova S, et al. Detection of hepatocellular carcinoma with PET/CT: a prospective comparison of 18F-fluorocholine and 18F-FDG in patients with cirrhosis or chronic liver disease. J Nucl Med. 2010 Nov;51(11):1699-706. PMID: 20956466.

288. Talbot J, Gutman F, Fartoux L, et al. PET/CT in patients with hepatocellular carcinoma using [(18)F]fluorocholine: preliminary comparison with [(18)F]FDG PET/CT. Eur J Nucl Med Mol Imaging. 2006 Nov;33(11):1285-9. PMID: 16802155.

289. Verhoef C, Valkema R, de Man RA, et al. Fluorine-18 FDG imaging in hepatocellular carcinoma using positron coincidence detection and single photon emission computed tomography. Liver. 2002 Feb;22(1):51-6. PMID: 11906619.

290. Wolfort RM, Papillion PW, Turnage RH, et al. Role of FDG-PET in the evaluation and staging of hepatocellular carcinoma with comparison of tumor size, AFP level, and histologic grade. Int Surg. 2010 Jan-Mar;95(1):67-75. PMID: 20480845.

291. Wu H-b, Wang Q-s, Li B-y, et al. F-18 FDG in conjunction with 11C-choline PET/CT in the diagnosis of hepatocellular carcinoma. Clin Nucl Med. 2011 Dec;36(12):1092-7. PMID: 22064078.

292. Wudel LJ, Jr., Delbeke D, Morris D, et al. The role of [18F]fluorodeoxyglucose positron emission tomography imaging in the evaluation of hepatocellular carcinoma. Am Surg. 2003 Feb;69(2):117-24; discussion 24-6. PMID: 12641351.

293. Yamamoto Y, Nishiyama Y, Kameyama R, et al. Detection of hepatocellular carcinoma using 11C-choline PET: Comparison with 18F-FDG PET. J Nucl Med. 2008 Aug;49(8):1245-8. PMID: 18632827.

294. Yoon KT, Kim JK, Kim DY, et al. Role of 18F-fluorodeoxyglucose positron emission tomography in detecting extrahepatic metastasis in pretreatment staging of hepatocellular carcinoma. Oncology. 2007;72 Suppl 1:104-10. PMID: 18087190.

295. Yoo HJ, Lee JM, Lee JY, et al. Additional value of SPIO-enhanced MR imaging for the noninvasive imaging diagnosis of hepatocellular carcinoma in cirrhotic liver. Investigative Radiology. 2009 Dec;44(12):800-7. PMID: 19838119.

296. Chen JG, Parkin DM, Chen QG, et al. Screening for liver cancer: results of a randomised controlled trial in Qidong, China. J Med Screen. 2003;10(4):204-9. PMID: 14738659.

297. Naaktgeboren CA, de Groot JA, van Smeden M, et al. Evaluating diagnostic accuracy in the face of multiple reference standards. Ann Intern Med. 2013 Aug 6;159(3):195-202. PMID: 23922065.

298. Hunt CH, Hartman RP, Hesley GK. Frequency and severity of adverse effects of iodinated and gadolinium contrast materials: retrospective review of 456,930 doses. AJR Am J Roentgenol. 2009 Oct;193(4):1124-7. PMID: 19770337.

299. Nelson KL, Gifford LM, Lauber-Huber C, et al. Clinical safety of gadopentetate dimeglumine. Radiology. 1995 Aug;196(2):439-43. PMID: 7617858.

300. Niendorf HP, Haustein J, Cornelius I, et al. Safety of gadolinium-DTPA: extended clinical experience. Magn Reson Med. 1991 Dec;22(2):222-8; discussion 9-32. PMID: 1812350.

301. Katayama H, Yamaguchi K, Kozuka T, et al. Adverse reactions to ionic and nonionic contrast media. A report from the Japanese Committee on the Safety of Contrast Media. Radiology. 1990 Jun;175(3):621-8. PMID: 2343107.

302. Micro-bubble Contrast Agents (marketed as Definity (Perflutren Lipid Microsphere) Injectable Suspension and Optison (Perflutren Protein-Type A Microspheres for Injection). Silver Spring, MD: U.S. Food and Drug Administration; 2008. http://www.fda.gov/safety/medwatch/safetyinformation/safetyalertsforhumanmedicalproducts/ucm092270.htm. Accessed on November 12, 2013.

303. Piscaglia F, Bolondi L. The safety of Sonovue in abdominal applications: retrospective analysis of 23188 investigations. Ultrasound Med Biol. 2006 Sep;32(9):1369-75. PMID: 16965977.

304. Herzog CA. Incidence of adverse events associated with use of perflutren contrast agents for echocardiography. JAMA. 2008 May 7;299(17):2023-5. PMID: 18460662.

305. Patient Safety: Radiation Dose in X-Ray and CT Exams. Oak Brook, IL: Radiological Society of North America; 2013. http://www.radiologyinfo.org/en/safety/index.cfm?pg=sfty_xray. Accessed on November 13, 2013.

306. Colli A, Fraquelli M, Casazza G, et al. Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. Am J Gastroenterol. 2006 Mar;101(3):513-23. PMID: 16542288.

307. Westwood M, Joore M, Grutters J, et al. Contrast-enhanced ultrasound using SonoVue(R) (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for

the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2013 Apr;17(16):1-243. PMID: 23611316.

308. Guang Y, Xie L, Ding H, et al. Diagnosis value of focal liver lesions with SonoVue-enhanced ultrasound compared with contrast-enhanced computed tomography and contrast-enhanced MRI: a meta-analysis. J Cancer Res Clin Oncol. 2011 Nov;137(11):1595-605. PMID: 21850382.

309. Xie L, Guang Y, Ding H, et al. Diagnostic value of contrast-enhanced ultrasound, computed tomography and magnetic resonance imaging for focal liver lesions: a meta-analysis. Ultrasound Med Biol. 2011 Jun;37(6):854-61. PMID: 21531500.

310. Chen L, Zhang L, Bao J, et al. Comparison of MRI with liver-specific contrast agents and multidetector row CT for the detection of hepatocellular carcinoma: a meta-analysis of 15 direct comparative studies. Gut. 2013 Oct;62(10):1520-1. PMID: 23929696.

311. Lin C-Y, Chen J-H, Liang J-A, et al. 18F-FDG PET or PET/CT for detecting extrahepatic metastases or recurrent hepatocellular carcinoma: a systematic review and meta-analysis. Eur J Radiol. 2012 Sep;81(9):2417-22. PMID: 21899970.

312. Aghoram R, Cai P, Dickinson JA. Alpha-foetoprotein and/or liver ultrasonography for screening of hepatocellular carcinoma in patients with chronic hepatitis B. Cochrane Database Syst Rev. 2012;9.

313. Lachs MS, Nachamkin I, Edelstein PH, et al. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. Ann Intern Med. 1992 Jul 15;117(2):135-40. PMID: 1605428.

314. Morana G, Salviato E, Guarise A. Contrast agents for hepatic MRI. Cancer Imaging. 2007;7 Spec No A:S24-7. PMID: 17921081.

315. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med. 2002 Jun 15;21(11):1525-37. PMID: 12111918.

316. Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. AJR Am J Roentgenol. 2006 Aug;187(2):271-81. PMID: 16861527.

317. Macaskill P, Gatsonis C, Deeks J, et al. Chapter 10. Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C, eds. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 1.0 ed. Oxford: The Cochrane Collaboration; 2010. Available from: http://srdta.cochrane.org/.

318. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005 Sep;58(9):882-93. PMID: 16085191.

319. Song F, Khan KS, Dinnes J, et al. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. Int J Epidemiol. 2002 Feb;31(1):88-95. PMID: 11914301.

Abbreviations and Acronyms

AASLD	American Association for the Study of Liver Diseases
AFP	Alpha-fetoprotein
AFROC	alternate free response operating characteristic
AHRQ	Agency for Healthcare Research and Quality
AUROC	area under the receiver operating characteristic
BCLC	Barcelona Clinic Liver Cancer
CER	Comparative effectiveness review
CT	Computed tomography
EPC	Evidence-based Practice Center
FDA	U.S. Food and Drug Administration
FDG	^{18}F-fluorodeoxyglucose
FLT	^{18} F-fluorothymidine
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
KQ	Key Question
LR	Likelihood ratio
MRI	Magnetic resonance imaging
PET	Positron emission tomography
PICOTS	Populations, interventions, comparators, outcomes, timing, setting
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
SIP	Scientific information packet
TEP	Technical expert panel
UCSF	University of California, San Francisco
US	Ultrasound